

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:49 ; Search time 65.7205 Seconds
(without alignments)
629.533 Million cell updates/sec

Title: US-10-026-001-1
Perfect score: 2120
Sequence: 1 MSFTTRSTFTSTNYRSLGSVQ.....RRIVDGKVVSETNDTKVLRH 430

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2120	100.0	430	2 S05481	keratin 18, type I
2	1808.5	85.3	423	2 I59463	keratin, type I, c
3	1137.5	53.7	368	2 A28825	keratin, type I no
4	917.5	43.3	472	1 KRHUE	keratin 14, type I
5	904.5	42.7	452	2 I49595	cytokekeratin 15 - m
6	891.5	42.1	473	2 JC4313	keratin 16, type I
7	889.5	42.0	399	2 A25470	cytokekeratin 19 - b
8	887	41.8	456	1 KRHUS	keratin 15, type I
9	880.5	41.5	403	2 JQ0028	cytokekeratin 19 - m
10	871.5	41.1	400	1 KRHU9	keratin 19, type I
11	870.5	41.1	467	2 I50476	keratin type I - g
12	865.5	40.8	432	2 S30433	keratin 17, type I
13	862.5	40.7	401	2 S57657	keratin 19 - potot
14	847.5	40.0	458	1 KRHU3	keratin 13, type I
15	846	39.9	437	2 A55682	keratin 13, type I
16	843.5	39.8	420	2 A37343	keratin 13, type I
17	839.5	39.6	593	1 KRHU0	keratin 10, type I
18	839	39.6	526	1 KRBOVI	keratin, 54K type
19	831	39.2	433	2 S01631	keratin, type I, e
20	817	38.5	569	1 KRMSE1	keratin, 59K type
21	815.5	38.5	561	2 A31994	keratin 10, type I
22	813.5	38.4	419	2 A25438	keratin, type I cy
23	813	38.3	424	2 S37780	keratin 20, type I
24	811	38.3	483	2 A55033	keratin 12 - mouse
25	810.5	38.2	359	2 B26135	keratin, 52K type
26	804	37.9	429	2 A25145	keratin, 47K type
27	800.5	37.8	486	1 KRXL	keratin 3, type I,
28	796.5	37.6	411	2 S45318	keratin 12 - rabbi
29	786.5	37.1	473	4 A33652	probable keratin 1

30	779.5	36.8	429	2 A40452	keratin 21, type I
31	769.5	36.3	570	2 S07330	keratin, epidermal
32	762.5	36.0	412	1 KRSHL1	keratin, 48K type
33	748.5	35.3	416	2 A61404	keratin A, type I
34	744.5	35.1	404	2 JS0073	keratin, 47.6K typ
35	740.5	34.9	416	2 S60034	keratin Hal, type
36	739.5	34.9	416	2 A46559	keratin, type I, h
37	725.5	34.2	392	2 A60777	keratin 2, type I,
38	722	34.1	407	2 I48739	MHa2(keratin acidi
39	712	33.6	362	2 I37459	keratin Ha3-II, ty
40	703	33.2	622	2 I37984	keratin 9, type I,
41	696.5	32.9	327	2 S04511	keratin 3, type I,
42	618.5	29.2	323	2 I48667	MHa3 (keratin acid
43	572	27.0	483	2 A34720	keratin 8, type II
44	570	26.9	466	2 A25074	vimentin - human
45	567	26.7	464	2 A56600	intermediate filam

ALIGNMENTS

RESULT 1

S05481.
keratin 18, type I, cytoskeletal - human
N/Alternate names: cytokekeratin 18
C/Species: Homo sapiens (man)
C/Date: 07-Sep-1990 #sequence revision 07-Sep-1990 #text_change 09-Jul-2004
C/Accession: S05481; S05482; S06889
R/Oshima, R.G.; Millan, J.L.; Cecena, G.
Differentiation 33, 61-68, 1986
A/Title: Comparison of mouse and human keratin 18: a component of intermediate filaments
A/Reference number: S05481; MUID:87134778; PMID:2434380
A/Accession: S05481
A/Molecule type: mRNA
A/Residues: 1-430 <OSH>
A/Cross-references: UNIPROT:P05783; EMBL:X12881; NID:G34036; PIDN:CAA31375.1; PID:G34037
R/Romano, V.; Hatzfeld, M.; Magin, T.M.; Zimbelmann, R.; Franke, W.W.; Maier, G.; Ponstir
Differentiation 30, 244-253, 1986
A/Title: Cytokekeratin expression in simple epithelia. I. Identification of mRNA coding for
A/Reference number: S05482; MUID:86193258; PMID:2422083
A/Accession: S05482
A/Molecule type: mRNA
A/Residues: 199-201, 'Q', 203-245, 'S', 247-308, 'R', 310-311, 'R', 313-430 <ROM>
A/Cross-references: EMBL:X12876; NID:G34034; PIDN:CAA31369.1; PID:G34035
A/Note: part of this sequence was confirmed by protein sequencing
R/Leube, R.E.; Bosch, F.X.; Romano, V.; Zimbelmann, R.; Hoefler, H.; Franke, W.W.
Differentiation 33, 69-85, 1986
A/Title: Cytokekeratin expression in simple epithelia.
A/Reference number: S06888; MUID:87134779; PMID:2434381
A/Accession: S06889
A/Molecule type: mRNA
A/Residues: 7-430 <LEU>
A/Cross-references: EMBL:X12883; NID:G30310; PIDN:CAA31377.1; PID:G30311
C/Genetics:
A/Gene: GDB:KRT18
A/Cross-references: GDB:120127; OMIM:148070
A/Map position: 17p12-17p11
C/Superfamily: cytoskeletal keratin
C/Keywords: coiled coil; intermediate filament
F/2-430/Product: keratin 18, type I, cytoskeletal #status predicted <MAT>

Query Match 100.0%; Score 2120; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.7e-96;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSFTTRSTFTSTNYRSLGSVQAPSYGARPVSSAASVYAGAGSGSRISVSRTSFRGGWGS	60
Db	1	MSFTTRSTFTSTNYRSLGSVQAPSYGARPVSSAASVYAGAGSGSRISVSRTSFRGGWGS	60
Qy	61	GGLATGIAGGLAGMGGIONEKETMQSLNDRLASLYLDRVSRLETENRRRLSKIREHLEKKG	120
Db	61	GGLATGIAGGLAGMGGIONEKETMQSLNDRLASLYLDRVSRLETENRRRLSKIREHLEKKG	120

```

QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
Db 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVEVKGLQQAQIASSGLTVEVDAP 240
Db 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVEVKGLQQAQIASSGLTVEVDAP 240
QY 241 KSQDLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTVVTQSAEVCAGAAETTLTEL 300
Db 241 KSQDLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTVVTQSAEVCAGAAETTLTEL 300
QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQQA 360
Db 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQQA 360
QY 361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRRIVDGKVS 420
Db 361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRRIVDGKVS 420
QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430

RESULT 2
I59463
keratin, type I, cytoskeletal - mouse
N:Alternate names: endo B cytokekeratin; keratin D
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 09-Jul-2004
C:Accession: I59463; A25621; A28428; JT0406
R:Alonso, A.; Weber, T.; Jorcano, J.L.
Roux's Arch. Dev. Biol. 196, 16-21, 1987
A>Title: Cloning and characterization of keratin D, a murine endodermal cytoskeletal pr
A:Reference number: I59463
A:Accession: I59463
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-423 <RES>
A:Cross-references: UNIPROT:P05784; GB:M36376; NID:G198587; PIDN:AAA39373.1; PID:G293682
R:Singer, P.A.; Trevor, K.; Oshima, R.G.
J. Biol. Chem. 261, 538-547, 1986
A>Title: Molecular cloning and characterization of the endo B cytokekeratin expressed in p
A:Reference number: A25621; MUID:86085876; PMID:2416755
A:Accession: A25621
A:Molecule type: mRNA
A:Residues: 1-243, 'D', 245-252, 'A', 254-423 <SIN>
A:Cross-references: GB:M11686; NID:G198620; PIDN:AAA39390.1; PID:G293685
R:Oshima, R.G.; Trevor, K.; Shevinsky, L.H.; Ryder, O.A.; Cecena, G.
Genes Dev. 2, 505-516, 1988
A>Title: Identification of the gene coding for the endo B murine cytokekeratin and its me
A:Reference number: A28428; MUID:88255838; PMID:2454868
A:Accession: A28428
A:Molecule type: DNA
A:Residues: 1-132 <OSH>
A:Cross-references: GB:Y00217; NID:G50842; PIDN:CAA68365.1; PID:G50843
R:Ichinose, Y.; Morita, T.; Zhang, F.; Srimalasongram, S.; Tondella, M.L.C.; Matsumoto,
Gene 70, 85-95, 1988
A>Title: Nucleotide sequence and structure of the mouse cytokekeratin endoB gene.
A:Reference number: JT0406; MUID:89196920; PMID:2467843
A:Accession: JT0406
A:Molecule type: DNA
A:Residues: 1-133, 'F', 135-243, 'D', 245-252, 'A', 254-423 <ICH>
A:Cross-references: GB:M22832; NID:G340757; PIDN:AAA37552.1; PID:G532610
C:Genetics:
A:Gene: endoB; KERG
A:Introns: 132/3; 160/2; 212/3; 267/3; 309/3; 384/2
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament

Query Match 85.3%; Score 1808.5; DB 2; Length: 423;
Best Local Similarity 85.6%; Pred. No. 2.5e-81;

```

```

Matches 370; Conservative 27; Mismatches 24; Indels 11; Gaps 4;
QY 1 MSFTTTS-TFSTNYRSLGSGVQAPSYGARFVPSAASVYAGAGSGSRISVSRSTSFGRGMG 59
Db 1 MSFTTSTTFSTNYRSLGSGVTPSQVRPAPSAASVYAGAGSGSRISVSRSV-WGGSVG 59
QY 60 SGLATGIAGGLAGGIONEKETMQSLNDRSLASLDYLRVRSLETENRRRLESKIREHLEKK 119
Db 60 S-----AGLAGGGIOTEKETMQDLNDRSLASLDYLRVRSLETENRRRLESKIREHLEKK 111
QY 120 GPQ-VRDWSHYFKIIEDLRAQIFANTVDNARIIVLQIDNARLAADDPRVKYETELAMRQSV 178
Db 112 GPQGVDRGWGHYFKIIEDLRAQILLANSVDNARIIVLQIDNARLAADDPRVKYETELAMRQSV 171
QY 179 ENDIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVEVKGLQQAQIASSGLTVEVD 238
Db 172 ESDIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVEVKGLQQAQIASSGLTVEVD 231
QY 239 APKSQDLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTVVTQSAEVCAGAAETTLTE 298
Db 232 APKSQDLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTVVTQSAEVCAGAAETTLTE 291
QY 299 LRTTQSLLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQR 358
Db 292 LRTTQSLLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQR 351
QY 359 QAQYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRRIVDGKV 418
Db 352 QAQYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRRIVDGKV 411
QY 419 VSETNDTKVLRH 430
Db 412 VSETNDTKVLRH 423

RESULT 3
A28825
keratin, type I nonepidermal - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A28825
R:Laflamme, S.E.; Jamrich, M.; Richter, K.; Sargent, T.D.; Dawid, I.B.
Genes Dev. 2, 853-862, 1988
A>Title: Xenopus endo B is a keratin preferentially expressed in the embryonic notochord.
A:Reference number: A28825; MUID:89092007; PMID:2463213
A:Accession: A28825
A:Molecule type: mRNA
A:Residues: 1-368 <LAF>
A:Cross-references: UNIPROT:P08802; GB:Y00230; NID:G64863; PIDN:CAA68372.1; PID:G64864
C:Genetics:
A:Start codon: GGT
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil

Query Match 53.7%; Score 1137.5; DB 2; Length 368;
Best Local Similarity 61.2%; Pred. No. 1e-48;
Matches 226; Conservative 74; Mismatches 58; Indels 11; Gaps 5;
QY 69 GGLAG-----MGGIONEKETMQSLNDRSLASLDYLRVRSLETENRRRLESKIREHLEKKGPQ 122
Db 2 GGFSGAGNVNLFGGVQNEKETMQDLNDRSLASLDYLRVRSLETENRRRLESKIREHLEKKGP- 60
QY 123 VRDWSHYFKIIEDLRAQIFANTVDNARIIVLQIDNARLAADDPRVKYETELAMRQSVENDI 182
Db 61 AKDWSPYMTIEDLKKQVFNISVENSQVLVQIDNARLAADDPRVKYETELAMRQSVENDI 120
QY 183 HGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVEVKGLQQAQIASSGLTVEVDAPKS 242
Db 121 GGLRLKVIDDTNITRLQLETEIEALKKEELLFMKKNHODDVLNQLQAQIATSAVTEVDAPKS 180
QY 243 QDLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTVVTQSAEVCAGAAETTLTELRT 302
Db 181 QDLGKIMADIRAOYDELRARKNREELDKYWSQOIEESTVVTQSAEVCAGAAETTLTELRT 240

```


Query Match 42.7%; Score 904.5; DB 2; Length 452;
Best Local Similarity 45.7%; Pred. No. 3e-37;
Matches 205; Conservative 77; Mismatches 124; Indels 43; Gaps 11;
QY 5 TRSTF---STNYSRL-----GSVQAPSY---GAPV---SSAASVYAGAGSGSRISVRS 51
Db TSSTFGSSTRGASLRAGGSFGGSLYGGGSRISASARFVSSGAGG---FGGMS 64
QY 52 TSFRGGSGGLATGIAGGLAGMG-----IQNEKETMQSLNDRSLASVLDVRSLETEN 105
Db CGFGGFG-GGFGGFGGDFGFGGGLLSGNEKVTMQLNDRSLASVLDKVRLEQAN 123
QY 106 RRLSKIREHLEKGPQV--RDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADD 163
Db 124 TELEKIRDWYQKQSPASPRDYSHYFKTMEIRDKILAAIDNSRVVLEIDNARLAADD 183
QY 164 FRVKEYETELAMQSVENDIHLKRVIDDNTNITRLQLETEIEALKEELLFMKNHEEEVKG 223
Db 184 FRLKYENELTLRQGVADINGLRLRVLDLTLARTDLEMQIEQLNEELAYLKNHEEEMKE 243
QY 224 LQAIASSGLTVEVDAPKQDLAKIMADIRAQYDELARKNREELDKYWSQIEESTTVVT 283
Db 244 FSSQLAGQ-VNVEDAAPGVLDLTMELAEQVYAEAKNRDVEAWFFSKTEELNKEVA 302
QY 284 TQSAEVAEATTLTELRRTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQLNGILL 343
Db 303 SNTMIQTSKTEITDLRRTLOGLLELQSLQSKAGLENSLAEEVCRYATLQIQIGVIT 362
QY 344 HLESELACTRAEGORQAQYEALNKKVLEAEATYRRLLED-----GEDFN 391
Db 363 GLETQLSELHSQEAQVQYNNMLLDIKTRLEQEIATYRNLLEDQDAKMAGIVREGSGG 422
QY 392 LGDALDSNSMQTIQKTTTRIVDGKVS 420
Db 423 GGSSSSSNFHISVEES-----VDGKVS 446

RESULT 6

JC4313
keratin 16, type I, cytoskeletal - human
N;Alternate names: 46k keratin type I; cytokeratin 16
C;Species: Homo sapiens (man)
C;Date: 06-Dec-1995 #sequence revision 08-Feb-1996 #text_change: 09-Jul-2004
C;Accession: JC4313; A24843; I58129
R;Paladini, R.D.; Takahashi, K.; Gant, T.M.; Coulombe, P.A.
Biochem. Biophys. Res. Commun. 215, 517-523, 1995
A;Title: cDNA cloning and bacterial expression of the human type I keratin 16.
A;Reference number: JC4313; MUID:96011809; PMID:7487986
A;Accession: JC4313
A;Molecule type: mRNA
A;Residues: 1-473 <PAL>
A;Cross-references: UNIPROT:P08779; GB:S79867; NID:G1195530; PIDN:AAB35421.1; PID:G11955
R;Experimental source: epidermal keratinocytes
R;RayChaudhury, A.; Marchuk, D.; Lindhurst, M.; Fuchs, E.
Mol. Cell. Biol. 6, 539-548, 1986
A;Title: Three tightly linked genes encoding human type I keratins: conservation of sequence
A;Reference number: A24843; MUID:87064338; PMID:2431270
A;Accession: A24843
A;Molecule type: DNA
A;Residues: 1, 'T', 3-25, 'A', 39-40, 43, 'ASTY', 48-49, 'A', 51-186, 'HAL', 190-207, 'ART'
A;Cross-references: GB:M28439; NID:G186683; PIDN:AAA59460.1; PID:G186685
A;Note: the translated sequence in GenBank entry HUMKER16A8, release 111.0, (PIDN:AAA594
R;McLean, W.H.I.; Rugg, E.L.; Lunny, D.P.; Morley, S.M.; Lane, E.B.; Swenson, O.; Doppi
Kunkeler, L.; Munro, C.S.
Nature Genet. 9, 273-278, 1995
A;Title: Keratin 16 and keratin 17 mutations cause pachyonychia congenita.
A;Reference number: I58129; MUID:95291318; PMID:7539673
A;Accession: I58129
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 118-131, 'P', 133-134 <MCL>
A;Cross-references: GB:S78514; NID:G1000376; PIDN:AAB34564.1; PID:G1000377
A;Note: this is a mutant sequence

C;Comment: This protein is an intermediate filament protein and expressed in epithelial t
ound healing, psoriasis and cancer.
C;Genetics:
A;Gene: GDB:KRT16
A;Cross-references: GDB:136207; OMIM:148067
A;Map position: 17pter-17qter
A;Note: defects in this gene may result in Jadassohn-Lewandowsky pachyonychia congenita
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 42.1%; Score 891.5; DB 2; Length 473;
Best Local Similarity 46.0%; Pred. No. 1.4e-36;
Matches 197; Conservative 82; Mismatches 122; Indels 27; Gaps 9;
QY 17 GSVQAPS-YG-----ARPVSSAASVYAGAGSGSRISVRSSTFRGGMGGGLATGIAG 69
Db 38 GSCRAPSTYGGGLSVSSRFSSGGACGLGGYGGGFSFSSSGFGGGYG-GGLGAGGG 96
QY 70 GL-AGMG-----IQNEKETMQSLNDRSLASVLDVRSLETERRRLESKIREHLEKK 119
Db 97 GLGAGFGGFGAGDGLLVGSEKVTMQLNDRSLASVLDKVRLEAEANADLEVKIRDWYQK 156
QY 120 GP-QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDLRAQVETELAMQSV 178
Db 157 RPSEIKDYSYFKTIEDLRNKIIAATENAQPIQIDNARLAADDLRAQVETELAMQSV 216
QY 179 ENDIHLKRVIDDNTNITRLQLETEIEALKEELLFMKNHEEEVKGLQAIASSGLTVEVD 238
Db 217 EADVNGLRRLVDELTLARTDLEMQIEGLKEELAYLRKNHEEEMALRGQ-TGGDVNVED 275
QY 239 APKSQLAKIMADIRAQYDELARKNREELDKYWSQIEESTTVVTTSARVAAETTLTE 298
Db 276 AAPGVLDLSRLNEMRDQYEQMAEKNRDAETWFLSKTEELNKEVASNSSELVQSSRSEVTE 335
QY 299 LRRTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGOR 358
Db 336 LRRVLQGLEIELOQSLSMKASLENSLEETKGRYCMQLSQIOGLIGSVVEQLAQLRCMEBQ 395
QY 359 QAQYEALNKKVLEAEATYRRLLEDGEDFNLDGALDSNSMQTIQKTTTRIVDGKV 418
Db 396 QSQEQYQLLDVKTRELEQEIATYRRLLE-GEDAHLE-----SSQASGQSISSEVFTSSS 449
QY 419 VSETNDTK 426
Db 449 SSSSRQTR 456

RESULT 7

A25470
cytokeratin 19 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C;Accession: A25470
R;Bader, B.L.; Magin, T.M.; Hatzfeld, M.; Franke, W.W.
EMBO J. 5, 1865-1875, 1986
A;Title: Amino acid sequence and gene organization of cytokeratin no. 19, an exceptional
A;Reference number: A25470; MUID:87004553; PMID:2428612
A;Accession: A25470
A;Molecule type: DNA
A;Residues: 1-399 <BAD>
A;Cross-references: UNIPROT:P08728; GB:X04152; NID:9469; PIDN:CRA27770.1; PID:G1197196
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 42.0%; Score 889.5; DB 2; Length 399;
Best Local Similarity 48.0%; Pred. No. 1.4e-36;
Matches 191; Conservative 80; Mismatches 114; Indels 13; Gaps 7;
QY 2 SFTTRSTFTNYSRLSGSVQAPS--YGARPVSSAASVYAGAGSGSRISVRSSTFRGGMG 59
Db 3 SYSYRQSSSTS--SFGMGGSGMRPFGAGAFRAPSIHGGSGRGVSVSSARFVS---SS 56
QY 60 SGGLATGIAGGLAGMGGI--QNEKETMQSLNDRSLASVLDVRSLETERRRLESKIREHLE 117

Db 57 SGGYGGYGALATSDGLLAGNEKLTWQNLDRSLASVLEKVRALBEANGDLEVKIRDWYQ 116
QY 118 KKGP-QVRDWSHYFKIIEDLRQAIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQ 176
Db 117 KQPGPARDYSHYFKTIEDLRDQILGATFENSIVLQIDNARLAADDFRTKFEQALRM 176
QY 177 SVENDIHGRKVIDDTNITRLQLETEIPALKEELLFMKNHHEEVKGLQAQIASSGLTVE 236
Db 177 SVEADINGLRVLDLTARTDLEMQIEGLKEELAYLKNHHEEEMSVLKGQVGGQ-VSVE 235
QY 237 VDAPKSDIAKIMADIRAQYDELARKNREELDKYWSQOIEESTTQSAEVAEATTL 296
Db 236 VDSAPGIDLAKILSDMRQYEVIAEKNRKDAEAWFISQTEELNREVAGHTEQLQISKTEV 295
QY 297 TELRRTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEG 356
Db 296 TDLRRTVQLEIELOQSLSMKAALEGYLAETEARFGAQLAQIQALISGIEAQLSDVRADT 355
QY 357 QROAQEYEAALLNIKVLEAEIATYRRLLLEDGEDFNLDG 394
Db 356 ERQNEYQYHLMKIDKTRLEQEIATYRNLLB-GQDAYFND 392

RESULT 8

KRHUS
keratin 15, type I, cytoskeletal - human
N;Alternate names: acidic cytokeatin; cytokeatin 15
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S01069; B61556; A3211; B30186
R;Leube, R.E.; Bader, B.L.; Bosch, F.X.; Zimbelmann, R.; Achtstaetter, T.; Franke, W.W.
J. Cell Biol. 106, 1249-1261, 1988
A;Title: Molecular characterization and expression of the stratification-related cytokeatin
A;Reference number: S01068; MUID:88198369; PMID:2452170
A;Accession: S01069
A;Molecule type: mRNA
A;Residues: 1-456 <LEU>
A;Cross-references: UNIPROT:P19012; EMBL:X07696; NID:G34070; PIDN:CAA30535.1; PID:G34071
R;Bader, B.L.; Jahn, L.; Franke, W.W.
Eur. J. Cell Biol. 47, 300-319, 1988
A;Title: Low level expression of cytokeatins 8, 18 and 19 in vascular smooth muscle cells containing the cytokeatin 19 gene.
A;Reference number: A61556; MUID:89210901; PMID:2468493
A;Accession: B61556
A;Molecule type: DNA
A;Residues: 1-166,195-456 <BAD>
A;Cross-references: GDB:120124; OMIM:148030
A;Map position: 17q21-17q22
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; intermediate filament
F;1-97/Domain: head <HEA>
F;98-413/Domain: helical rod #status predicted <ROD>
F;414-456/Domain: tail <TAI>

Query Match 41.8%; Score 887; DB 1; Length 456;
Best Local Similarity 44.2%; Pred. No. 2.2e-36;
Matches 199; Conservative 85; Mismatches 126; Indels 40; Gaps 10;

QY 4 TTRSTF---STNYSRL-----GSVQAPSYGARPVSSAASVYAGAGGSGSRISVSRST 52
Db 8 TSSSTFGGSTRGGSLLAGGGGFGGSLSGGGSGSRISASSARFVSSGGGGYGGGMVRC 67
QY 53 SFRGGMGs---GGLATGIAGGL-----AGMGGI---QNEKETMQLNDRSLASVLDVRSL 102
Db 68 GFGGGAGSVFGGGFGGGFGGGGGGGLSGNEKITMQLNDRSLASVLDKVRAL 127
QY 103 TENRRLESKIREHLEKGPQVR--DWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLA 160

Db 128 EANADLEVKIHWYQKQTPASPECDSYQFKTIEELRDKIMATTIDNSRVILEIDNARLA 187
QY 161 ADDRVRKYETELAMRQSVENDIHGRKVIDDTNITRLQLETEIEALKKEELLFMKNHHEE 220
Db 188 ADDRKLKYENELALRQGVADINGLRRLVLDLTARTDLEMQIEGLNEELAYLKNHHEE 247
QY 221 VKGLQAQIASSGLTVEVDAPKSDIAKIMADIRAQYDELARKNREELDKYWSQOIEESTT 280
Db 248 MKPESSQLAGQ-VNVEMDAAPGVLDLTVLAEMREQYEAEMAEKNRRDVEAWFFSKTEELNK 306
QY 281 VVTQSAEVAEATTLTELRTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQLNG 340
Db 307 EVASNTEMIQTETITDLRRTMQLIELOQSLSMKAGLENSLAETECRYATQLQIQG 366
QY 341 ILLHLESELAQTRAEGQRAQYEAALLNIKVLEAEIATYRRLLLEDGEDFNLDG 392
Db 367 LIGGLEAQLSELRCMEAEQYKMLLDIKTRLEQEIATYRSLLB-GQDAKMGAGIGIREA 425
QY 393 --GDALDSSNSMOTIQKTTTTRIVDGVKVS 420
Db 426 SSGGGSSSNFHNVEES-----VDGQVVS 450

RESULT 9

JQ0028
cytokeatin 19 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1990 #sequence revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: JQ0028; JQ0378; JN0101
R;Ichinose, Y.; Hashido, K.; Miyamoto, H.; Nagata, T.; Nozaki, M.; Morita, T.; Matsushir
Gene 80, 315-323, 1989
A;Title: Molecular cloning and characterization of cDNA encoding mouse cytokeatin No. 1
A;Reference number: JQ0028; MUID:90060780; PMID:2479589
A;Accession: JQ0028
A;Molecule type: mRNA
A;Residues: 1-403 <ICH>
A;Cross-references: UNIPROT:P19001; GB:M28698; NID:G233167; PIDN:AAA60432.1; PID:G6233168
R;Lussier, M.; Quellet, T.; Lampron, C.; Lapointe, L.; Royal, A.
Gene 85, 435-444, 1989
A;Title: Mouse keratin 19: complete amino acid sequence and gene expression during devel
A;Reference number: JQ0378; MUID:90185218; PMID:2483396
A;Accession: JQ0378
A;Molecule type: mRNA
A;Residues: 1-403 <LUS>
R;Lussier, M.; Fillon, M.; Compton, J.G.; Nadeau, J.H.; Lapointe, L.; Royal, A.
Gene 95, 203-213, 1990
A;Title: The mouse keratin 19-encoding gene: sequence, structure and chromosomal assignme
A;Reference number: JN0101; MUID:91065533; PMID:1701153
A;Accession: JN0101
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <LUS2>
A;Cross-references: GB:M36120; NID:G198583; PIDN:AAA39371.1; PID:G387393
C;Comment: Cytokeatin polypeptides are major components of intermediate filaments which
C;Genetics:
A;Gene: K19
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil
F;1-75/Domain: head <HED>
F;76-118/Domain: coil 1a, alpha-helical rod <C1A>
F;133-228/Domain: coil 1b, alpha-helical rod <C1B>
F;249-390/Domain: coil 2, alpha-helical rod <HC2>

Query Match 41.5%; Score 880.5; DB 2; Length 403;
Best Local Similarity 50.1%; Pred. No. 3.9e-36;
Matches 191; Conservative 78; Mismatches 93; Indels 19; Gaps 8;

QY 17 GSVQAPSYGARPVSSAASVYAGAGGSGSRISVSR-----STSP---RGMGSGGLATGIA 68
Db 21 GSVRISSGG---VFRAPSIHGGSGGRGVSVSTRFVTSSSGSYGVGRGSGFSGTLA--VS 75
QY 69 GGLAGMGGIQNEKETMQLNDRSLASVLDVRSLTENRRLESKIREHLEKGP-QVRDWS 127

QY 85 QSLNDRLASLYLDRVRSLETENRRLSEKIREHLEKK-GPOVRDWSHYPKIIEDLRQAIFAN 143
Db 121 QNLNDRLATYLEKVHSLEKANGDLIELKIRQFLENKTSPDARDYSAYHATISDLQDMIQDA 180

QY 144 TVDNARIVQLIDNAPLAADDPRVKYETELAMRQSVENDIHGLRKVIDDTNITRLQLETEI 203
Db 181 TRINGGVYLAI DNAKLATDDFKTKYENELAMRQSVEADIAGLKRLDELTLARSLEMOI 240

QY 204 EALKEELLFMKKHHEEVKGLOAQIASGCLTVEVDPAKSQDLAKIMADIRAOYDELRKN 263
Db 241 EGLKEELIYLLKXNHEEELASMRQMTGT-VNVEVDAAPQEDLSRVMAHIREQYEGVS AKN 299

QY 264 REELDKYWSQQIEBSTTVTTQSAEVLGAEEAETLTTELRRVTQSQSLEIDLDSMRNLKASLENS 323
Db 300 ORELDAAWFQTSETILTKEVTANTETTLQVSKTEVTTELRRTIQGLEIELQSELSKKRSLEGT 359

QY 324 LREVEARYALQMEQNGILLHLESELAAQTRAEGQRQAQYEALLNINIKVLEAEATYRRL 383
Db 360 LATDESRYSIQLTOLQARVTSLEEIQVHLRGMDRQSQBYQMLLDINKTRLEMEIAEYRRL 419

QY 384 LEDG-EDFNL-----GDALSSNSMQTTQKTTTTRRIVDGKVVSET 422
Db 420 LDGGATSFSTSGGGGGGGVSVSTKITTV-KTIEEDIVDGKVVST 465

RESULT 12
S30433
N:Alternate names: cytoskeletal - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S30433; S23648; I78538; B29139
R:Troyanovsky, S.M.; Leube, R.E.; Franke, W.W.
Eur. J. Cell Biol. 59, 127-137, 1992
A>Title: Characterization of the human gene encoding cytokeratin 17 and its expression pattern
A:Reference number: S30433; MUID:93105967; PMID:1281771
A:Accession: S30433
A:Molecule type: DNA
A:Residues: 1-432 <TRO>
A:Cross-references: UNIPROT:Q04695; EMBL:Z19574; NID:g30378; PIDN:CAA79626.1; PID:g30379
R:Flohr, T.; Buwitt, U.; Bonnekoh, B.; Decker, T.; Boettger, E.C.
Eur. J. Immunol. 22, 975-979, 1992
A>Title: Interferon-gamma regulates expression of a novel keratin class I gene.
A:Reference number: S23648; MUID:92201318; PMID:1372562
A:Accession: S23648
A:Molecule type: mRNA
A:Residues: 1-432 <FLO>
A:Cross-references: EMBL:X62571; NID:g34074; PIDN:CAA44451.1; PID:g34075
R:McLean, W.H.I.; Rugg, E.L.; Lunny, D.P.; Morley, S.M.; Lane, E.B.; Swennesson, O.; Doppelt,
Kunkeler, L.; Munro, C.S.
Nature Genet. 9, 273-278, 1995
A>Title: Keratin 16 and keratin 17 mutations cause pachyonychia congenita.
A:Reference number: I58129; MUID:95291318; PMID:7539673
A:Accession: I78538
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 85-91, 'D', 93-101 <MCL>
A:Cross-references: GB:S78515; NID:gi000378; PIDN:AAB34565.1; PID:gi000379
R:Kartasova, T.; Cornelissen, B.J.C.; Belt, P.; van de Putte, P.
Nucleic Acids Res. 15, 5945-5962, 1987
A>Title: Effects of UV, 4-NQO and TPA on gene expression in cultured human epidermal keratinocytes
A:Reference number: A29139; MUID:87316861; PMID:2442723
A:Accession: B29139
A:Molecule type: mRNA
A:Residues: 167, 'E', 169-175, 'D', 177-432 <KAR>
A:Cross-references: GB:X05803; NID:g34080; PIDN:CAA29248.1; PID:g34081
C:Genetics:
A:Gene: GDB:KRT17
A:Cross-references: GDB:I36211; OMIM:148069
A:Map position: 17q12-17q21
A:Introns: 144/3; 172/2; 224/3; 278/3; 320/3; 394/2; 402/1
A>Note: defects in this gene may result in Jackson-Lawler pachyonychia congenita
C:Superfamily: cytoskeletal keratin

```
C;Keywords: coiled coil

Query Match      40.8%; Score 865.5; DB 2; Length 432;
Best Local Similarity 44.6%; Pred. No. 2.3e-35;
Matches 189; Conservative 83; Mismatches 115; Indels 37; Gaps 8;

QY 27 RPVSSAASV--YAGAGGSGSRISVSRSSTSFRRGGMGSLATGIAGGL----- 71
DB 6 RQFTSSSIKIGSSGLGGSSRTSCLRLS---GGLGAGSCRLGAGGLSTLGGSSSYSCV 61
QY 72 -----AGVGGIQ-----NEKETMQSLNDRLASLYLDVRVSRLETENRRLESKIREHLE 117
DB 62 SFGSGGYGSFPGVDGLLAGGEKATQNINLRNLASYLDKVRALAEANTELEVKIRDWYQ 121
QY 118 KKGP-QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQ 176
DB 122 RQAPGPARDYSQYRTIEELQNKILATVDNANILLQIDNARLAADDPRTKFETEQAURL 181
QY 177 SVENDIHGLRKVIDDTNIRLOLETEIHALKEELLFMKNHEEVKGLQAIASSGLTVE 236
DB 182 SVEADINGLRRVLDELTLARADLEMQIENLKEELAYLKHNHEEMNALRGQVGGE-INVE 240
QY 237 VDAPKSODLAKIMADIRAQYDELARKOREELDXYKSQQIESTVTVTTSQAEVGAETTL 296
DB 241 MDAPGVDSLRIINEMRDQYEKMAEKNRKDAEDWFFSKTEELNREVATNSLVQSGKSEI 300
QY 297 TELRRTVQSLEIDLDSMRNLKASLENSLRVEARYALQMEQLNGILLHLESELAQTRAEG 356
DB 301 SELRRTMQALETELQSQLSMKASLEGFLAETENRYCVQLSQIQGLIGSVBEQLAQRCM 360
QY 357 QRQAQEYEAALLNIKVKLEABIAFYRRLLLEDGEDFNLGDALDSSNSMQTIQKTTRRIVDG 416
DB 361 EQNQEQYKILLDVKRTRLEQEIATYRRULE-GEDAHLTYKKPEVTTROV-RTIVEEVQDG 418
QY 417 KVVVS 420
DB 419 KVVIS 422.

RESULT 13
S57657
keratin 19 - potoroo
C;Species: Potorous tridactylus (potoroo, long-nosed rat kangaroo)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S66257; S57657
R;Boettger, V.; Stasiak, P.C.; Harrison, D.L.; Mellerick, D.M.; Lane, E.B.
Eur. J. Biochem. 231, 475-485, 1995
A;Title: Epitope mapping of monoclonal antibodies to keratin 19 using keratin fragments,
A;Reference number: S66257; MUID:95361872; PMID:7543411
A;Accession: S66257
A;Molecule type: mRNA
A;Residues: 1-401 <BOE>
A;Cross-references: UNIPROT:P51856; EMBL:X82579; NID:g886903; PIDN:CAAS7915.1; PID:g8869
C;Genetics:
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; intermediate filament
F;1-73/Domain: head <HEA>
F;74-386/Domain: helical rod #status predicted <ROD>

Query Match      40.7%; Score 862.5; DB 2; Length 401;
Best Local Similarity 46.6%; Pred. No. 2.9e-35;
Matches 183; Conservative 82; Mismatches 119; Indels 9; Gaps 6;

QY 2 SFTRTFSTNYRSLGVSQAPSYGARPVSSAAVYACAGGSGSRISVSRSTSFRGGMGSG 61
DB 3 SYSYRQSSSTLTGLGSNNAAAFGS-GAFRAPSHGGSGGLGVSVSRFASSGLSGGYGG 61
QY 62 GLATGIAGGLAGMGGI--ONEKETMQSLNDRLASLYLDVRVSRLETENRRLESKIREHLEKH 119
DB 62 --SSSFVGYGGADGLLAGNEKITMQNINLRNLASYLDKVRALAEANADLEVKIRDWCQXQ 119
QY 120 GP-QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSV 178
```

```

Db      120 GPGPARDYSAYMTTQDLRDKILGATIENSKIVLQIDNARLAADDFRTKFEETEARLSV 179
QY      179 ENDIHGRKVIDDITNITRLQLEIEALKEELFMKXNHEBEVKGLOAQIASSGLTVEVD 238
Db      180 EADINGLRRVDELTLARADLEMEQLEGLKEELAYLKKHHEEISALSQGVGGQ-VSVEVD 238
QY      239 APKSQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVTTTQSAEVGAETTLTE 298
Db      239 SAPGIDLAKILTDMRSQYEAAMVEKNRSDAEAWFTSKTDELNOEAVHTKLLQTSKTEVTD 298
QY      299 LRRVTQSLRIDLDNRNLKASLENSLRVEARYALQMEQLNGILLHLESELAQTRAEGQR 358
Db      299 LRRTLQGLEIELOSLQSMKAALLEGTLAETEARXGVQLSQIALISNIESQLSDIRADMER 358
QY      359 QAOEYEAALLNIKVKLEAEIATYRRLLE--DGED 389
Db      359 QNQEYLLMDIKSLREREIATYRSLLEGHDAQD 391

```

RESULT 14

KRUH3

keratin 13, type I, cytoskeletal, long splice form - human

N:Alternate names: cyto keratin 13

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 09-Jul-2004

C:Accession: S06088; A33216; B37343; A33403

R:Mischke, D.; Wachter, E.; Hochstrasser, K.; Wild, A.G.; Schulz, P.

Nucleic Acids Res. 17, 7984, 1989

A:Title: The N-, but not the C-terminal domains of human keratins 13 and 15 are closely

A:Reference number: S06088; MUID:90016882; PMID:2477803

A:Accession: S06088

A:Molecule type: mRNA

A:Residues: 1-458 <NIS1>

A:Cross-references: UNIPROT:P13646; EMBL:X14640; NID:G34032; PIDN:CAA32786.1; PID:G34033

A:Accession: A33216

A:Molecule type: protein

A:Residues: 291-299 <MIS2>

R:Kuruc, N.; Leube, R.E.; Moll, I.; Bader, B.L.; Franke, W.W.

Differentiation 42, 111-123, 1989

A:Title: Synthesis of cyto keratin 13, a component characteristic of internal stratified

A:Reference number: A37343; MUID:90228645; PMID:2483837

A:Accession: B37343

A:Molecule type: mRNA

A:Residues: 1-57, 'G', 59-458 <KUR>

A:Cross-references: GB:X52426

R:Schulz, P.; Wachter, E.; Hochstrasser, K.; Wild, A.G.; Mischke, D.

Biochem. Biophys. Res. Commun. 162, 1522-1527, 1989

A:Title: Sequence of a human keratin 13 specific cDNA encompassing coil 1B through the 3

A:Reference number: A33403; MUID:89350978; PMID:2475110

A:Accession: A33403

A:Molecule type: mRNA

A:Residues: 158-458 <SCH>

C:Genetics:

A:Gene: GDB:KRT13

A:Cross-references: GDB:120740; OMIM:148065

A:Map position: 17q21-17q22

C:Superfamily: cytoskeletal keratin

C:Keywords: alternative splicing; coiled coil; intermediate filament

F:1-96/Domain: head <HEA>

F:97-412/Domain: helical rod #status predicted <ROD>

F:413-458/Domain: tail <TAI>

Query Match

Best Local Similarity 40.0%; Score 847.5; DB 1; Length 458;

Matches 192; Conservative 88; Mismatches 139; Indels 37; Gaps 10;

QY

1 MSFTTRSTFTSTNYRSLGVSQAPSYGARPVSSAASVY---ACAGGSGSRISV---SRSTS 53

Db

1 MSRLQSSASVGGGFGGSCQLGGGRGVSTCTRFVSGSAGGYGGVSCGFGGGADSG 60

QY

54 FRGGWG-----SGCLATGIAGLAG-----MGGI--QNEKETMOSLNDRLASLYLD 96

```

Db      61 FGGYGGGLGGYGGGLGGGFGGAGGFVDFGACDGLLTGNEKITMQLNDRLASYLE 120
QY      97 RVRSLFETNRRLSKIREHLEKKGK--QVRDWSHYFKIIEDIRAQIFANTVDNARIVLQI 154
Db      121 KVRALBANADLEVKIRDWHLKQSPASPERDYSPYKTIIELRDKILTATTENNRVILEI 180
QY      155 DNARLAADDFRVKYETELAMRQSVENDIHGLRKVIDDTNITRLQLETEIEALKEELFMK 214
Db      181 DNARLAVDDFRLKYENELALRQSVADINGLRRVLDDELTLSTKDLQMIEISLNEELAYMK 240
QY      215 KNHEEEVKGLQQAIASSGLTVEVDAPKSDQLAKIMADIRAOYDELARKNREELDKYWSQ 274
Db      241 KNHEEEMKEFSNQVVQ--VNVENDATPGIDLTVLAEMREQYEAMAEERNRDAEWFHAK 299
QY      275 IEESTTVVTTQSAEVGAETTLTLELRRTVQSLRIDLDNRNLKASLENSLRVEARYALQ 334
Db      300 SAELNKEVSTNTAMITQTSKTEITELRRTLQGLEIELOSLQSMKAALLEGTLAETEARXGVQLSQIALISNIESQLSDIRADMER 359
QY      335 MEQLNGILLHLESELAQTRAEGQRQAQYEAALLNIKVKLEAEIATYRRLLE--DGED 392
Db      360 LQQIQGLISSIEAQLSELRSMECQNEQYKMLLDIKTRLEQBIATYRSLLE-GQDAKMIG 418
QY      393 -----GDALDSSNSMQTIQ--KTTTRRIVDGKVYSE 421
Db      419 FPSSAGSVSPRSTSVTTTSSASVTTSNASGRRTSD 454

```

RESULT 15

A55682

keratin 13, type I cytoskeletal - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-Mar-1995 #sequence revision 03-Mar-1995 #text_change 09-Jul-2004

C:Accession: A55682; B23518; A60493

R:Fillon, M.; Sarafian, V.; Lussier, M.; Belanger, C.; Lapointe, L.; Royal, A.

Genomics 24, 303-310, 1994

A:Title: Arrangement of a cluster of three mouse type I keratin genes expressed sequentia

A:Reference number: A55682; MUID:95213021; PMID:7535287

A:Accession: A55682

A:Molecule type: DNA

A:Residues: 1-437 <FIL>

A:Cross-references: UNIPROT:P08730; GB:U13921; NID:G541610; PIDN:AAC52150.1; PID:G541611

R:Knapp, B.; Rentrop, M.; Schweizer, J.; Winter, H.

Nucleic Acids Res. 14, 751-763, 1986

A:Title: Nonepidermal members of the keratin multigene family: cDNA sequences and in situ

A:Reference number: A93640; MUID:86120369; PMID:2418416

A:Accession: B23518

A:Molecule type: mRNA

A:Residues: 126-437 <KNA>

A:Cross-references: GB:X03492; NID:G52782; PIDN:CAA27208.1; PID:G52783

R:Winter, H.; Rentrop, M.; Nischt, R.; Schweizer, J.

Differentiation 43, 105-114, 1990

A:Title: Tissue-specific expression of murine keratin K13 in internal stratified squamous

lation state of a distinct Cpg site in the remote 5'-flanking region of the gene.

A:Reference number: A60493; MUID:90323435; PMID:1695590

A:Accession: A60493

A:Molecule type: DNA

A:Residues: 1-157 <WIN>

A:Cross-references: GB:X53320; NID:G288245; PIDN:CAA37407.1; PID:G288246

C:Genetics:

A:Gene: K13

A:Introns: 157/3; 185/2; 237/3; 291/3; 333/3; 407/2; 415/1

C:Superfamily: cytoskeletal keratin

C:Keywords: coiled coil; intermediate filament

Query Match

Best Local Similarity 39.9%; Score 846; DB 2; Length 437;

Matches 187; Conservative 73; Mismatches 113; Indels 36; Gaps 9;

QY

25 GARPVSSAASVYAGAGSGSRISVSRSTSFRCGMGSGGLATGIAGLAG----- 73

Db

25 GGRNISSCSRFRV-TGGSAGGYGGGWSGCGFGGAG-GGFGGFGGFGGFGGFGG 82

QY

74 -MGGIQ-----NEXETMQSLNDRLASLYLDRVSRLETENRRRLSKIREHLEKKGK--QVR 124

Db 83 DFGVGGLLSGNEKITMQNLNDRLASYLDKVRALANADLEVKIRDWHLKQSPASPER 142
Qy 125 DWSHFYKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVENDING 184
Db 143 DYSAYYKTIEELRIKILEATTNNRIILEIDNARLAADDFRLKYENELTLRQSVREADING 202
Qy 185 LRKVIDDTNITRLQLETEIEALKEELLFMKNHEEEVKGLOAQIASSGLTVEVDAPKSQD 244
Db 203 LRRVDELTLAKTDLEMQIESLNEELAYLKKNHEEEMKEFNSQVVQ-VNVEMDATPGID 261
Qy 245 LAKIMADIRAQYDELAQRNREELDKYWSQOIEESTVVTQSAEVGAAETTLTELRRTVQ 304
Db 262 LTRVLAEMREQYEALAEKNRRDAEEWFQTKSAELNKEVSSNAEMIQTSTKTEITELRRTLQ 321
Qy 305 SLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQORQAQEYE 364
Db 322 GLEIELQSLSMKAGLESTLAETECRYALQLOQIQIQLISSIEAQLSELRSMECQNEQYK 381
Qy 365 ALNLIKVKLEAEIATYRRLLIEDGED-----FNLG-----DALDSSNS 401
Db 382 MLLDIKTRLEQEIATYRSLLE-GQDAKWTGFNSGCGNNTTTSNGSPSSNS 429

Search completed: December 14, 2004, 09:20:18
Job time : 67.7205 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31 ; Search time 331.419 Seconds
(without alignments)
465.433 Million cell updates/sec

Title: US-10-026-001-1

Perfect score: 2120

Sequence: 1 MSFTTRSTFTSTNYSLSVQ.....RRIVDGKVVSEINDTKVLRH 430

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2120	100.0	430	4 AAB90795	Aab90795 Human she
2	2120	100.0	430	5 ABB81087	Abb81087 Human cyt
3	2120	100.0	430	5 ABP68606	Abp68606 Human pan
4	2120	100.0	430	8 ADE76940	Ad76940 Human pro
5	2120	100.0	430	8 ADF91429	Adf91429 Human cyt
6	2120	100.0	447	4 ABG16550	Abg16550 Novel hum
7	2120	100.0	452	4 AAG74328	Aag74328 Human col
8	2120	100.0	454	8 ADP04217	Adp04217 Human col
9	2115	99.8	429	6 AAE38083	Aae38083 Human cyt
10	1996.5	94.2	409	6 ABO14647	Abol14647 Novel hum
11	1995.5	94.1	456	4 ABG15224	Abg15224 Novel hum
12	1980	93.4	449	4 ABG08564	Abg08564 Novel hum
13	1858	87.6	378	6 ABO14648	Abol14648 Novel hum
14	1817.5	85.7	423	8 ADF91430	Adf91430 Mouse cyt
15	1792	84.5	675	4 ABG14106	Abg14106 Novel hum
16	1770	83.5	359	6 AAE38087	Aae38087 Human cyt
17	1725	81.4	428	4 ABG08563	Abg08563 Novel hum
18	1708.5	80.6	427	4 AAU10061	Aau10061 Cytokerat
19	1708.5	80.6	427	7 ADJ87755	Adj87755 G-coupled
20	1708.5	80.6	427	8 ADI79311	Adi79311 NOV4 prot
21	1708.5	80.6	427	8 ADO55990	Ado55990 Human NOV
22	1655	78.1	475	4 ABG17250	Abg17250 Novel hum
23	1627	76.7	433	4 ABG16694	Abg16694 Novel hum
24	1598.5	75.4	428	4 ABG23266	Abg23266 Novel hum
25	1550.5	73.1	478	4 AAU30550	Aau30550 Novel hum

26	1548	73.0	399	4	ABG14471	Abg14471 Novel hum
27	1545.5	72.9	417	6	AAE33677	Aae33677 Human str
28	1472	69.4	364	4	ABG27717	Abg27717 Novel hum
29	1467.5	69.2	326	4	ABG15279	Abg15279 Novel hum
30	1460	68.9	382	4	ABG15222	Abg15222 Novel hum
31	1362	64.2	627	7	ADJ70961	Adj70961 Human hea
32	1293	61.0	360	6	ABU52621	Abu52621 Human NOV
33	1266	59.7	332	4	ABG18693	Abg18693 Novel hum
34	1240.5	58.5	302	4	ABG17249	Abg17249 Novel hum
35	1223	57.7	298	4	ABG18630	Abg18630 Novel hum
36	1215.5	57.3	298	4	ABG01045	Abg01045 Novel hum
37	1156	54.5	267	8	ADP04216	Adp04216 Human col
38	1152.5	54.4	321	4	ABG00672	Abg00672 Novel hum
39	1137	53.6	513	4	ABG16693	Abg16693 Novel hum
40	1124	53.0	310	4	ABG09381	Abg09381 Novel hum
41	1121.5	52.9	296	8	ADP04215	Adp04215 Human col
42	1117	52.7	315	8	ADP04214	Adp04214 Human col
43	1110	52.4	232	5	AAU84288	Aau84288 Human end
44	1088	51.3	545	4	ABG18629	Abg18629 Novel hum
45	1048.5	49.5	385	4	ABG27716	Abg27716 Novel hum

ALIGNMENTS

RESULT 1
AAB90795
ID AAB90795 standard; protein; 430 AA.
XX
AC AAB90795;
XX
DT 15-JUN-2001 (first entry)
XX
DE Human shear stress-response protein SEQ ID NO: 90.
XX
KW Human; shear stress-response protein; vascular disease; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO200125427-A1.
XX
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000WO-JP006840.
XX
PR 01-OCT-1999; 99JP-00280976.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (NOJI/) NOJIMA H.
XX
PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
PI Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
DR WPI; 2001-266308/27.
DR N-PSDB; AAH02918.
XX
PT DNA sequences, proteins encoded by them and antibodies against them
PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis.
XX
PS Claim 60; Page 482-484; 678pp; Japanese.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension
XX
SQ Sequence 430 AA;

Query Match 100.0%; Score 2120; DB 4; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.2e-141;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAASVYAGAGSGSRISVSSTFRGGMGS 60
 DB 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAASVYAGAGSGSRISVSSTFRGGMGS 60
 QY 61 GGLATGIAGGLAGMGGIQTNEKTMQSLNDRSLASVYDRVSRLETENRRLESKIREHLEKKG 120
 DB 61 GGLATGIAGGLAGMGGIQTNEKTMQSLNDRSLASVYDRVSRLETENRRLESKIREHLEKKG 120
 QY 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
 DB 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
 QY 181 DIHGLRKVIDDNTNITRLOLETEIEALKEBELLFMKNHHEEVKGLQAOIASGLTVEVDAP 240
 DB 181 DIHGLRKVIDDNTNITRLOLETEIEALKEBELLFMKNHHEEVKGLQAOIASGLTVEVDAP 240
 QY 241 KSQDLAKIMADIRAOYDELRARKEBELDKYWSQIQIEESTTQSAEVGAETTLTEL 300
 DB 241 KSQDLAKIMADIRAOYDELRARKEBELDKYWSQIQIEESTTQSAEVGAETTLTEL 300
 QY 301 RTVQSLDLDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
 DB 301 RTVQSLDLDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
 QY 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVV 420
 DB 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVV 420
 QY 421 ETNDTKVLRH 430
 DB 421 ETNDTKVLRH 430

RESULT 2

ABP68606
 ID ABP68606 standard; protein; 430 AA.
 XX
 AC ABP68606;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human cytokeratin 18 (CK18) protein.
 XX
 KW Antigen; cytokeratin 18; CK-18; immune response; cytostatic; vaccine;
 KW gene therapy; cancer; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200255555-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 21-DEC-2001; 2001WO-US049964.
 XX
 PR 21-DEC-2000; 2000US-0257820P.
 XX
 PA (GENZ) GENZYME CORP.
 XX
 PI Nicolette CA;
 XX
 DR WPI; 2002-619103/66.
 DR N-PSDB; ABN86610.
 XX
 PT Novel antigenic cytokeratin 18 compounds and peptides useful for inducing
 PT an immune response in a subject and for diagnosing a neoplastic condition
 PT or susceptibility to the condition of an animal cell or tissue.
 XX
 PS Claim 6; Page 71-72; 73pp; English.
 XX
 CC The invention relates to novel antigenic cytokeratin 18 (CK-18) compounds
 CC (I) and peptides useful for inducing an immune response in a subject. (I)
 CC is useful for inducing an immune response in a subject, by delivering (I)

CC as a polynucleotide, in the context of an MHC molecule that presents the
 CC compound on the surface of an APC. (I) is useful for aiding in the
 CC diagnosis of the neoplastic condition or susceptibility to the condition
 CC of an animal cell or tissue; for generating antibodies which are useful
 CC for identifying and purifying polypeptides and APCs expressing the
 CC polypeptides. (I) serves as markers for the neoplastic phenotype. (I)
 CC that is covalently or non-covalently linked to molecules are useful in
 CC diagnostic methods, and for detecting or purifying antibodies. It is also
 CC useful as components of anti-cancer vaccines and to expand immune
 CC effector cells that are specific for cells having differential, i.e.
 CC aberrant, expression of antigenic CK-18. CK-18 proteins or antibodies are
 CC useful for detecting, diagnosing or prognosing and monitoring the
 CC progression, course or stage of CK-18 related cancers, or malignancies.
 CC Host cells comprising one or more immunogenic ligands are useful for
 CC inducing an immune response in a subject, and to expand a population of
 CC immune effector cells such as tumour infiltrating lymphocytes which in
 CC turn are useful in adoptive immunotherapies. Agents that modulate the
 CC binding of CK-18 protein to its ligand are useful for treating disease,
 CC especially cancer. The present sequence represents the human CK18 protein
 XX
 SQ Sequence 430 AA;

Query Match 100.0%; Score 2120; DB 5; Length 430;
 Best Local Similarity 100.0%; Pred. No. 1.2e-141;

Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAASVYAGAGSGSRISVSSTFRGGMGS 60
 DB 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAASVYAGAGSGSRISVSSTFRGGMGS 60
 QY 61 GGLATGIAGGLAGMGGIQTNEKTMQSLNDRSLASVYDRVSRLETENRRLESKIREHLEKKG 120
 DB 61 GGLATGIAGGLAGMGGIQTNEKTMQSLNDRSLASVYDRVSRLETENRRLESKIREHLEKKG 120
 QY 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
 DB 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
 QY 181 DIHGLRKVIDDNTNITRLOLETEIEALKEBELLFMKNHHEEVKGLQAOIASGLTVEVDAP 240
 DB 181 DIHGLRKVIDDNTNITRLOLETEIEALKEBELLFMKNHHEEVKGLQAOIASGLTVEVDAP 240
 QY 241 KSQDLAKIMADIRAOYDELRARKEBELDKYWSQIQIEESTTQSAEVGAETTLTEL 300
 DB 241 KSQDLAKIMADIRAOYDELRARKEBELDKYWSQIQIEESTTQSAEVGAETTLTEL 300
 QY 301 RTVQSLDLDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
 DB 301 RTVQSLDLDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
 QY 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVV 420
 DB 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVV 420
 QY 421 ETNDTKVLRH 430
 DB 421 ETNDTKVLRH 430

RESULT 3

ABP68606
 ID ABP68606 standard; protein; 430 AA.
 XX
 AC ABP68606;
 XX
 DT 14-JAN-2003 (first entry)
 XX
 DE Human pancreatic cancer expressed protein SEQ ID NO 155.
 XX
 KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
 KW cytostatic; tumour.
 XX
 OS Homo sapiens.

XX WO2002060317-A2.
XX 08-AUG-2002.
XX 30-JAN-2002; 2002WO-US002781.
XX 30-JAN-2001; 2001US-0265305P.
XX 31-JAN-2001; 2001US-0265682P.
XX 09-FEB-2001; 2001US-0267568P.
XX 21-MAR-2001; 2001US-0278651P.
XX 28-APR-2001; 2001US-0287112P.
XX 16-MAY-2001; 2001US-0291631P.
XX 12-JUL-2001; 2001US-0305484P.
XX 20-AUG-2001; 2001US-0313999P.
XX 27-NOV-2001; 2001US-0333626P.
XX (CORI-) CORIXA CORP.
XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
PI WPI; 2002-627435/67.
PI N-PSDB; ABV94751.
XX New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT diagnosing, preventing and/or treating cancer, particularly pancreatic
PT cancer.
XX Claim 2; SEQ ID NO 155; 300pp + Sequence Listing; English.
XX The invention relates to an isolated polynucleotide (I) comprising: (a)
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC complements of (a); (c) sequences consisting of at least 20 contiguous
CC residues of (a); (d) sequences that hybridize to (a), under moderately
CC stringent conditions; (e) sequences having at least 75% or 90% identity
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
CC in a patient and compositions comprising polypeptides, polynucleotides,
CC antibodies, fusion proteins, T cell populations and antigen presenting
CC cells expressing the polypeptide are useful in treating pancreatic cancer
CC and stimulating an immune response. The polynucleotides can be used as
CC probes or primers for nucleic acid hybridization, in the design and
CC preparation of ribozyme molecules for inhibiting expression of the tumor
CC polypeptides and proteins in the tumor cells, in vaccines and for gene
CC therapy. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 430 AA;
SQ

Query Match 100.0%; Score 2120; DB 5; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.2e-141;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFTTRSTSTNYSRSLGVSQAPSYGARPVSAAVSAGAGGSGRISVSRSSTFRGGMGS 60
Db 1 MSFTTRSTSTNYSRSLGVSQAPSYGARPVSAAVSAGAGGSGRISVSRSSTFRGGMGS 60
QY 61 GGLATGIAGGLAGMGGIONEKETMQSLNDRLASVLDVRSLETENRLESKIREHLEKKG 120
Db 61 GGLATGIAGGLAGMGGIONEKETMQSLNDRLASVLDVRSLETENRLESKIREHLEKKG 120
QY 121 PQVRDWSHVFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFVKYETELAMRQSVEN 180
Db 121 PQVRDWSHVFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFVKYETELAMRQSVEN 180
QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKBELLFMKKNHEEVKGLQAOIASSGLTVEVDAP 240
Db 181 DIHGLRKVIDDNTNITRLQLETEIEALKBELLFMKKNHEEVKGLQAOIASSGLTVEVDAP 240
QY 241 KSQDLAKIMADIRAOYDELAKNREELDKYWSQIIEESTTWTTSQAEVGAETTLTLR 300
Db 241 KSQDLAKIMADIRAOYDELAKNREELDKYWSQIIEESTTWTTSQAEVGAETTLTLR 300

QY 301 RTVQSLIEDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEQORQA 360
Db 301 RTVQSLIEDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEQORQA 360
QY 361 QEYEAALLNIKVKLEAEIATYRLLLEDGEDFNLDSSNSMQTIQKTTTRRIVDGKVV 420
Db 361 QEYEAALLNIKVKLEAEIATYRLLLEDGEDFNLDSSNSMQTIQKTTTRRIVDGKVV 420
QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430
RESULT 4
ADE76940
ID ADE76940 standard; protein; 430 AA.
XX ADE76940;
AC ADE76940;
XX 29-JAN-2004 (first entry)
XX Human protein expressed in a liver disorder #30.
DE human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
XX Homo sapiens.
OS US2003108871-A1.
XX 12-JUN-2003.
XX 30-JUL-2001; 2001US-00919039.
XX 28-JUL-2000; 2000US-0222113P.
XX (KASE/) KASER M R.
XX Kaser MR;
XX WPI; 2004-031227/03.
XX N-PSDB; ADE76939.
XX Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX Claim 1; SEQ ID NO 105; 41pp; English.
XX The invention relates to a composition comprising several cDNAs that are
CC differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.
XX Sequence 430 AA;
SQ

Query Match 100.0%; Score 2120; DB 8; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.2e-141;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAAVYAGAGGSGSRISVSSTSFRRGGMGS 60
DB 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAAVYAGAGGSGSRISVSSTSFRRGGMGS 60

QY 61 GGLATGIAGGLAGMGGIQNEKETMOSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 120
DB 61 GGLATGIAGGLAGMGGIQNEKETMOSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 120

QY 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
DB 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180

QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 240
DB 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 240

QY 241 KSODLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTTTVTTSQAEVGAETTLTLELR 300
DB 241 KSODLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTTTVTTSQAEVGAETTLTLELR 300

QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQROA 360
DB 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQROA 360

QY 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRIVDGKVVVS 420
DB 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRIVDGKVVVS 420

QY 421 ETNDTKVLRH 430
DB 421 ETNDTKVLRH 430

RESULT 5
ADF91429
ID ADF91429 standard; protein; 430 AA.
XX ADF91429;
AC ADF91429;
XX 26-FEB-2004 (first entry)
XX Human cytokerin 18 #SEQ ID 1.
DE Antiasthmatic; antiallergic; antiinflammatory; gene therapy;
KW bronchial asthma; chronic rhinitis; cytokerin 18.
XX Homo sapiens.
XX WO2003098211-A1.
XX 27-NOV-2003.
XX 12-MAY-2003; 2003WO-KR000933.
XX 15-MAY-2002; 2002KR-00026765.
XX (NAHM/) NAHM D.
XX (JEON/) JEON S.
XX Nahm D, Jeon S;
XX WPI; 2004-022911/02.
XX Diagnosing bronchial asthma and chronic rhinitis comprises detecting
PT autoantibodies to cytokerin 18 protein in the bodily fluid of a human
PT subject.
XX Claim 8; SEQ ID NO 1; 42pp; English.

CC The invention relates to a method for diagnosing bronchial asthma and
CC chronic rhinitis comprising detecting autoantibodies to cytokerin 18
CC protein in the bodily fluid of a human subject. The method or cytokerin
CC 18 protein is useful in diagnosing, classifying bronchial asthma and
CC chronic rhinitis or for formulating pharmaceutical formulations for
CC protecting or treating patients with or non-allergic patients with
CC bronchial asthma and chronic rhinitis or patients with bronchial asthma
CC and chronic rhinitis associated with autoantibodies to cytokerin 18.
CC Cytokeratin 18 protein is used as drug target in the production of drugs
CC for treating bronchial asthma and chronic rhinitis. The current sequence
CC represents human cytokerin 18.
XX
SQ Sequence 430 AA;

Query Match 100.0%; Score 2120; DB 8; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.2e-141;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAAVYAGAGGSGSRISVSSTSFRRGGMGS 60
DB 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARPVSSAAVYAGAGGSGSRISVSSTSFRRGGMGS 60

QY 61 GGLATGIAGGLAGMGGIQNEKETMOSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 120
DB 61 GGLATGIAGGLAGMGGIQNEKETMOSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 120

QY 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
DB 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180

QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 240
DB 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 240

QY 241 KSODLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTTTVTTSQAEVGAETTLTLELR 300
DB 241 KSODLAKIMADIRAOYDELRARKNREELDKYWSQOIEESTTTVTTSQAEVGAETTLTLELR 300

QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQROA 360
DB 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQROA 360

QY 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRIVDGKVVVS 420
DB 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRIVDGKVVVS 420

QY 421 ETNDTKVLRH 430
DB 421 ETNDTKVLRH 430

RESULT 6
ABG16550
ID ABG16550 standard; protein; 447 AA.
XX ABG16550;
AC ABG16550;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #16541.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS80737.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 46909; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 447 AA;

Query Match 100.0%; Score 2120; DB 4; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.3e-141;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFTTRSTFTNYSRLSGVQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTSFSGMGMS 60
DB 18 MSFTTRSTFTNYSRLSGVQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTSFSGMGMS 77
QY 61 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRLESKIREHLEKKG 120
DB 78 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRLESKIREHLEKKG 137
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVEN 180
DB 138 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVEN 197
QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVKGLQAIASSGLTVEVDAP 240
DB 198 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVKGLQAIASSGLTVEVDAP 257
QY 241 KSQDLAKIMADIRAQYDELAARKNREELDKYWSQQIEESTVVTQSAEVAETTLTEL 300
DB 258 KSQDLAKIMADIRAQYDELAARKNREELDKYWSQQIEESTVVTQSAEVAETTLTEL 317
QY 301 RTVQSLEIDLDSNRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEQROA 360
DB 318 RTVQSLEIDLDSNRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEQROA 377
QY 361 QVEYALLNKKVLEAEIATYRLELLEDGEDFNGLDLSNSMQTIQKTTTRRIVDGKVS 420
DB 378 QVEYALLNKKVLEAEIATYRLELLEDGEDFNGLDLSNSMQTIQKTTTRRIVDGKVS 437
QY 421 ETNDTKVLRH 430

DB 438 ETNDTKVLRH 447
RESULT 7
AAG74328
ID AAG74328 standard; protein; 452 AA.
XX AC AAG74328;
XX DT 03-SEP-2001 (first entry)
XX DE Human colon cancer antigen protein SEQ ID NO:5092.
XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX KW colorectal carcinoma; chromosome 12.
XX OS Homo sapiens.
XX PN WO200122920-A2.
XX PD 05-APR-2001.
XX PF 28-SEP-2000; 2000WO-US026524.
XX PR 29-SEP-1999; 99US-0157137P.
XX PR 03-NOV-1999; 99US-0163280P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX DR WPI; 2001-235357/24.
XX DR N-PSDB; AAH33759.
XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX PS Claim 11; Page 6811-6813; 9803pp; English.
XX CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
XX CC proteins are collectively known as colon cancer antigens. The colon
XX CC cancer antigens have cytostatic activity and can be used in gene therapy
XX CC and vaccine production. N and P may be used in the prevention, diagnosis
XX CC and treatment of diseases associated with inappropriate P expression. For
XX CC example, N and P may be used to treat disorders associated with decreased
XX CC expression by rectifying mutations or deletions in a patient's genome
XX CC that affect the activity of P by expressing inactive proteins or to
XX CC supplement the patients own production of P. Additionally, N may be used
XX CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
XX CC into a host cell and culturing the cell to express the proteins. N and P
XX CC can be used in the prevention, diagnosis and treatment of colorectal
XX CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
XX CC sequences used in the exemplification of the present invention. N.B.
XX CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
XX CC time of publication, meaning no sequences are present for SEQ ID NO:1027
XX CC to 1052, 7921 and 7922
SQ Sequence 452 AA;
Query Match 100.0%; Score 2120; DB 4; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.3e-141;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFTTRSTFTNYSRLSGVQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTSFSGMGMS 60
DB 23 MSFTTRSTFTNYSRLSGVQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTSFSGMGMS 82
QY 61 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRLESKIREHLEKKG 120
DB 83 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRLESKIREHLEKKG 142

QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVEN 180
Db 143 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVEN 202
QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 240
Db 203 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 262
QY 241 KSQDLAKIMADIRAOYDELRKRNREELDKYWSQIIEESTTVVTTQSAEVGAAETTLTEL 300
Db 263 KSQDLAKIMADIRAOYDELRKRNREELDKYWSQIIEESTTVVTTQSAEVGAAETTLTEL 322
QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGORQA 360
Db 323 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGORQA 382
QY 361 QEYEALLNIKVKLEAEIATYRRLLLEDGEDFNLDGALDSSNSMOTIQKTTTTRIVDGKVV 420
Db 383 QEYEALLNIKVKLEAEIATYRRLLLEDGEDFNLDGALDSSNSMOTIQKTTTTRIVDGKVV 442
QY 421 ETNDTKVLRH 430
Db 443 ETNDTKVLRH 452

RESULT 8
ADP04217
ID ADP04217 standard; protein; 454 AA.

XX AC ADP04217;
XX DT 09-SEP-2004 (first entry)
XX DE Human colon specific protein SEQ ID NO:198.
XX KW human; colon specific nucleic acid; CSNA; colon specific protein; CSP;
XX KW cytostatic; vaccine; gene therapy; colon cancer.
XX OS Homo sapiens.

XX PN W02004050858-A2.
XX PD 17-JUN-2004.
XX PF 04-DEC-2003; 2003WO-US038808.
XX PR 04-DEC-2002; 2002US-0431133P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Macina RA, Turner LR, Sun Y, Rodriguez M, Burcham TS;
XX WPI; 2004-480622/45.

XX PT Novel colon specific protein derived from normal and neoplastic colon
cell, useful as vaccine in treating colon cancer and in identifying,
PT diagnosing, monitoring, staging, imaging colon cancer and non-cancerous
PT disease state in colon.

XX PS Claim 12; SEQ ID NO 198; 655pp; English.

XX CC The invention relates to a novel colon specific protein (CSP) (I), and
the nucleic acid encoding it. A CSP of the invention has cytostatic
activity, and may have use in a vaccine, and in gene therapy. The CSP is
useful for determining the presence of a colon specific protein in a
sample. The nucleic acid encoding the CSP is useful for determining the
presence of a colon specific nucleic acid (CSNA) in a sample. The CSP and
CSNA are useful for diagnosing or monitoring the presence and metastases
of colon cancer in a patient. The method of administering a composition
comprising a CSP or CSNA is useful for treating a patient with colon
cancer. The CSP and CSNA are useful as vaccine for treating colon cancer
and non-cancerous disease states in colon. The present sequence
represents a CSP of the invention.

XX SQ Sequence 454 AA;
Query Match 100.0%; Score 2120; DB 8; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.3e-141;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTSFRCGWS 60
Db 25 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTSFRCGWS 84
QY 61 GGLATGIAGGLAGMGCIQNEKETNQSLNDRLASYLDRVSRLETERRRLESKIREHLEKKG 120
Db 85 GGLATGIAGGLAGMGCIQNEKETNQSLNDRLASYLDRVSRLETERRRLESKIREHLEKKG 144
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVEN 180
Db 145 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVEN 204
QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 240
Db 205 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLOAQIASSGLTVEVDAP 264
QY 241 KSQDLAKIMADIRAOYDELRKRNREELDKYWSQIIEESTTVVTTQSAEVGAAETTLTEL 300
Db 265 KSQDLAKIMADIRAOYDELRKRNREELDKYWSQIIEESTTVVTTQSAEVGAAETTLTEL 324
QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGORQA 360
Db 325 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGORQA 384
QY 361 QEYEALLNIKVKLEAEIATYRRLLLEDGEDFNLDGALDSSNSMOTIQKTTTTRIVDGKVV 420
Db 385 QEYEALLNIKVKLEAEIATYRRLLLEDGEDFNLDGALDSSNSMOTIQKTTTTRIVDGKVV 444
QY 421 ETNDTKVLRH 430
Db 445 ETNDTKVLRH 454

RESULT 9
AAE38083
ID AAE38083 standard; protein; 429 AA.

XX AC AAE38083;
XX DT 06-NOV-2003 (first entry)
XX DE Human cytokeratin K18 protein.
XX KW Human; cancer-associated epitope; cytokeratin K8; cytokeratin K18;
KW adenocarcinoma; therapy; cancer.

XX OS Homo sapiens.

XX PN W02003057168-A2.

XX PD 17-JUL-2003.

XX PF 03-JAN-2003; 2003WO-US000297.

XX PR 03-JAN-2002; 2002US-0345208P.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Ditzel H, Jensenius JC;

XX WPI; 2003-598315/56.

XX DR N-PSDB; AAD57400.

XX PT Novel isolated cancer-associated epitope comprising two separate
polypeptides, a cytokeratin 8 polypeptide and a cytokeratin 18
polypeptide, useful as component of vaccine for preventing or treating

PT adenocarcinoma.
XX
PS Disclosure; Page 19; 155pp; English.
XX
CC The invention provides a cancer-associated epitope comprising two
CC separate polypeptides, a cytokeratin 8 polypeptide and a cytokeratin 18
CC polypeptide. Vaccine composition of the invention is useful for treating
CC or preventing colon adenocarcinoma, ovarian adenocarcinoma, renal
CC adenocarcinoma, mammary adenocarcinoma, lung adenocarcinoma, pancreatic
CC adenocarcinoma or non-seminoma testis carcinoma. The invention is also
CC useful for preparing a medicament for treating or preventing cancer in a
CC mammal. The present sequence is human cytokeratin K18 protein
XX
SQ Sequence 429 AA;

Query Match 99.8%; Score 2115; DB 6; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.7e-141;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SFTTRSTFSTNYSRLGSGVQAPSYGARPVSSAASVYAGAGGSGRSISVSRSTSFRCGMGSG 61
Db 1 SFTTRSTFSTNYSRLGSGVQAPSYGARPVSSAASVYAGAGGSGRSISVSRSTSFRCGMGSG 60
QY 62 GLATGIAGGLAGMGGQIQNEKTMQSLNDRILASLYLDRVRSLETENRRLSKIREHLEKKG 121
Db 61 GLATGIAGGLAGMGGQIQNEKTMQSLNDRILASLYLDRVRSLETENRRLSKIREHLEKKG 120
QY 122 QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRSQVEND 181
Db 121 QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRSQVEND 180
QY 182 IHGLRKVIDDNTNITRLQLETEIEALKEELFMKQNEHEBEVKGLOAQIASSGLTVEVDAPK 241
Db 181 IHGLRKVIDDNTNITRLQLETEIEALKEELFMKQNEHEBEVKGLOAQIASSGLTVEVDAPK 240
QY 242 SQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTTSQAEVGAETTLTELRR 301
Db 241 SQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTTSQAEVGAETTLTELRR 300
QY 302 TVQSLDIDLSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAEQORQAQ 361
Db 301 TVQSLDIDLSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAEQORQAQ 360
QY 362 EYEALLNIKVKLEAEIATYRLLLEDGEDENLGDALDSSNSMOTIQKTTTTRRIVDGKVVSE 421
Db 361 EYEALLNIKVKLEAEIATYRLLLEDGEDENLGDALDSSNSMOTIQKTTTTRRIVDGKVVSE 420
QY 422 TNDTKVLRH 430
Db 421 TNDTKVLRH 429

RESULT 10
ABO14647
ID ABO14647 standard; protein; 409 AA.
XX
AC ABO14647;
XX
DT 25-AUG-2003 (first entry)
XX
DE Novel human protein #20.
XX
KW Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
stroke; infection.

OS Homo sapiens.
XX
PN WO2003023002-A2.
XX
PD 20-MAR-2003.
XX
PF 09-SEP-2002; 2002WO-US028539.
XX
PR 07-SEP-2001; 2001US-0318120P.
PR 07-SEP-2001; 2001US-0318130P.
PR 10-SEP-2001; 2001US-0318430P.
PR 17-SEP-2001; 2001US-0322636P.
PR 17-SEP-2001; 2001US-0322781P.
PR 17-SEP-2001; 2001US-0322816P.
PR 17-SEP-2001; 2001US-0322817P.
PR 19-SEP-2001; 2001US-0323519P.
PR 20-SEP-2001; 2001US-0323631P.
PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325091P.
PR 26-SEP-2001; 2001US-0324990P.
PR 17-APR-2002; 2002US-0373212P.
PR 06-SEP-2002; 2002US-00236177.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
DR WPI; 2003-313242/30.
DR N-PSDB; ACD19340.
XX
PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT stroke or infections.
XX
PS Claim 1; Page 128-129; 586pp; English.
XX
CC The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This is the amino acid sequence of a novel human NOV protein
XX
SQ Sequence 409 AA;

Query Match 94.2%; Score 1996.5; DB 6; Length 409;
Best Local Similarity 95.1%; Pred. No. 6.3e-133;
Matches 409; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 MSFTTRSTFSTNYSRLGSGVQAPSYGARPVSSAASVYAGAGGSGRSISVSRSTSFRCGMGSG 60
Db 1 MSFTTRSTFSTNYSRLGSGVQAPSYGARPVSSAASVYAGAGGSGRSISVSRSTSFRCGMGSG 60

QY 61 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 120
 DB 40 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 99
 QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRSVEN 180
 DB 100 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRSVEN 159
 QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEVKGLQAQIASGLTVEVDAP 240
 DB 160 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEVKGLQAQIASGLTVEVDAP 219
 QY 241 KSQDLAKIMADIRAOYDELARKNREELDKYWSQQIEESTTVVTTQSAEVGAETTLTEL 300
 DB 220 KSQDLAKIMADIRAOYDELARKNREELDKYWSQQIEESTTVVTTQSAEVGAETTLTEL 279
 QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQQA 360
 DB 280 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQQA 339
 QY 361 QEYEALLNKKVLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRIVDGKVS 420
 DB 340 QEYEALLNKKVLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRIVDGKVS 399
 QY 421 ETNDTKVLRH 430
 DB 400 ETNDTKVLRH 409

RESULT 11

ABG15224
 ID ABG15224 standard; protein; 456 AA.

XX AC ABG15224;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #15215.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS79411.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 45583; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 456 AA;

Query Match 94.1%; Score 1995.5; DB 4; Length 456;

Best Local Similarity 94.5%; Pred. No. 8.5e-133;

Matches 409; Conservative 9; Mismatches 12; Indels 3; Gaps 2;

QY 1 MSFTTRSTFSTNYRSLGVSQAPSPVSSAASVYAGAGSGSRISVSRSTSRFGMGWS 60

DB 24 MSFTTRSTFSTNYRSLGVSQAPSPVSSAASVYAGAGSGSRISVSRSTSRFGMGWS 83

QY 61 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 120

DB 84 GGLATGIAGGLAGMGGQIQNEKETMQSLNDRSLASVLDVRSLETENRRLESKIREHLEKKG 143

QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDN--ARIVLQIDNARLAADDPRVKYETELA-NRQS 177

DB 144 PQVRDWSHYFKIIEDLRAQIFANTVVEQCPTTVLLRIDSRARLAADDPRVKYETEAAGPCAQS 203

QY 178 VENDIHLGRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEVKGLQAQIASGLTVEV 237

DB 204 VENNIHGLCKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEVKGLQAQIASGLTVEV 263

QY 238 DAPKSQDLAKIMADIRAOYDELARKNREELDKYWSQQIEESTTVVTTQSAEVGAETTLT 297

DB 264 DAPKSQDLAKIMADIRAOYDELARKNREELDKYWSQQIEESTTVVTTQSAEVGAETTLT 323

QY 298 ELRRTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQ 357

DB 324 ELRRTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQ 383

QY 358 RQAQYEALLNKKVLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRIVDGK 417

DB 384 RQAQYEALLNKKVLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRIVDGK 443

QY 418 VVSETNDTKVLRH 430

DB 444 VVSETNDTKVLRH 456

RESULT 12

ABG08564

ID ABG08564 standard; protein; 449 AA.

XX AC ABG08564;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8555.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS72751.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 38923; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABC30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 449 AA;

Query Match 93.4%; Score 1980; DB 4; Length 449;
Best Local Similarity 95.2%; Pred. No. 1e-131;
Matches 413; Conservative 2; Mismatches 15; Indels 4; Gaps 4;

QY 1 MSFTRSTFTNTRSLGVSQAPSYGARPVSSAASVYAGAGGSGRSISVSRSTSFRCGMGS 60
Db |||||
61 GGLATGIAGGLAGGQIENEKQSLNDRLASVLDVRSLETENRRLKSKIREHLEKKG 120
Db |||||
76 GGLATGIAGGLAGGQIENEKQSLNDVASYLDVRSLETENRRLKSKIREHLEKKG 135
QY 121 PQVDRWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVXYETELAMRQSVEN 180
Db |||||
136 PQVDRWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVXYETELAMRQSVEN 195
QY 181 DIHGLRKVIDDTNITRLQLETEALKEELFMKKNHEEVKGLQAIASSGLTVEVDAP 240
Db |||||
196 DIHGLRKVIDDTNITRLQLETEALKEELFMKKNHEEVKGLQAIASSGLTVEVDAP 255
QY 241 KSQDLAKIMADIRAQYDELA-RKNREELDKYWSQIEESTVTTTQSAEVGAETTLTEL 299
Db |||||
256 KSQDLAKIMADIRAQYXRAYGLKXNREELDKYWSQIEESTVTTTQSAEVGAETTLTEL 315
QY 300 RRTVQSLEIDLDNRNLKASLENSLREVEARYA-LQMEQL-NGILLHSESELAQTRAEGQ 357
Db |||||
316 RRTVQSLEIDLDNRNLKASLENSLREVEARTPLQMEQLQGRGSLHSESELAQTRAEGQ 375
QY 358 RQAEYEALLNIK-VKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTRIVDG 416

Db 376 RQAEYEALLLEHQSSLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTRIVDG 435
QY 417 KVVSETNDTKVLRH 430
Db 436 KVVSETNDTKVLRH 449

RESULT 13
ABO14648
ID ABO14648 standard; protein; 378 AA.
XX
AC ABO14648;
XX
DT 25-AUG-2003 (first entry)
XX
DE Novel human protein #21.
XX
KW Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection.
XX
OS Homo sapiens.
XX
PN WO2003023002-A2.
XX
PD 20-MAR-2003.
XX
XX 09-SEP-2002; 2002WO-US028539.
XX 07-SEP-2001; 2001US-0318120P.
XX 07-SEP-2001; 2001US-0318130P.
XX 10-SEP-2001; 2001US-0318430P.
XX 17-SEP-2001; 2001US-0322636P.
XX 17-SEP-2001; 2001US-0322781P.
XX 17-SEP-2001; 2001US-0322816P.
XX 17-SEP-2001; 2001US-0322817P.
XX 19-SEP-2001; 2001US-0323519P.
XX 20-SEP-2001; 2001US-0323631P.
XX 20-SEP-2001; 2001US-0323636P.
XX 25-SEP-2001; 2001US-0324969P.
XX 25-SEP-2001; 2001US-0325091P.
XX 26-SEP-2001; 2001US-0324990P.
XX 17-APR-2002; 2002US-0373212P.
XX 06-SEP-2002; 2002US-00236177.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
DR WPI; 2003-313242/30.
DR N-PSDB; ACD19341.
XX
XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT stroke or infections.
XX
PS Claim 1; Page 129; 586pp; English.
XX
CC The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,

CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This is the amino acid sequence of a novel human NOV protein
XX
SQ Sequence 378 AA;

Query Match 87.6%; Score 1858; DB 6; Length 378;
Best Local Similarity 100.0%; Pred. No. 3.6e-123; Indels 0; Gaps 0;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARVPVSSAASVYAGAGGSGSRISVSRSTFRGGMGS 60
DB 1 MSFTTRSTFSTNYRSLGSGVQAPSYGARVPVSSAASVYAGAGGSGSRISVSRSTFRGGMGS 60
QY 61 GGLATGIAGGLAGMGGIONEKETMQSLNDRSLASVLDNRSLRLETKIREHLEKKG 120
DB 61 GGLATGIAGGLAGMGGIONEKETMQSLNDRSLASVLDNRSLRLETKIREHLEKKG 120
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
DB 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKEELLFMKKNHEEVKGLQAIASSGLTVEVDAP 240
DB 181 DIHGLRKVIDDNTNITRLQLETEIEALKEELLFMKKNHEEVKGLQAIASSGLTVEVDAP 240
QY 241 KSQDLAKIMADIRAQYDELRARKNREELDKYWSQQIEESTVTTVTTQSAEVAETTLTEL 300
DB 241 KSQDLAKIMADIRAQYDELRARKNREELDKYWSQQIEESTVTTVTTQSAEVAETTLTEL 300
QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
DB 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
QY 361 QEYEALLNIKVKLEAEIA 378
DB 361 QEYEALLNIKVKLEAEIA 378

RESULT 14
ADF91430
ID ADF91430 standard; protein; 423 AA.

AC ADF91430;

XX 26-FEB-2004 (first entry)

XX Mouse cytokerin 18 #SEQ ID 2.

XX Antiasthmatic; antiallergic; antiinflammatory; gene therapy;
KW bronchial asthma; chronic rhinitis; cytokerin 18.

OS Mus sp.

XX WO2003098211-A1.

XX 27-NOV-2003.

XX

PF 12-MAY-2003; 2003WO-KR000933.
XX
PR 15-MAY-2002; 2002KR-00026765.
XX
PA (NAHM/) NAHM D.
PA (JEON/) JEON S.
XX
PI Nahm D, Jeon S;
XX
DR WPI; 2004-022911/02.
XX
XX Diagnosing bronchial asthma and chronic rhinitis comprises detecting
PT autoantibodies to cytokerin 18 protein in the bodily fluid of a human
PT subject.
XX
PS Claim 18; SEQ ID NO 2; 42pp; English.
XX
CC The invention relates to a method for diagnosing bronchial asthma and
CC chronic rhinitis comprising detecting autoantibodies to cytokerin 18
CC protein in the bodily fluid of a human subject. The method or cytokerin
CC 18 protein is useful in diagnosing, classifying bronchial asthma and
CC chronic rhinitis or for formulating pharmaceutical formulations for
CC protecting or treating patients with or non-allergic patients with
CC bronchial asthma and chronic rhinitis or patients with bronchial asthma
CC and chronic rhinitis associated with autoantibodies to cytokerin 18.
CC Cytokeratin 18 protein is used as drug target in the production of drugs
CC for treating bronchial asthma and chronic rhinitis. The current sequence
CC represents mouse cytokerin 18.
XX
SQ Sequence 423 AA;

Query Match 85.7%; Score 1817.5; DB 8; Length 423;
Best Local Similarity 86.1%; Pred. No. 3e-120;
Matches 372; Conservative 26; Mismatches 23; Indels 11; Gaps 4;

QY 1 MSFTTRS-TFSTNYRSLGSGVQAPSYGARVPVSSAASVYAGAGGSGSRISVSRSTFRGGMG 59
DB 1 MSFTTRS-TFSTNYRSLGSGVQAPSYGARVPVSSAASVYAGAGGSGSRISVSRSV-WGGSVG 59
QY 60 SGGLATGIAGGLAGMGGIONEKETMQSLNDRSLASVLDNRSLRLETKIREHLEKKG 119
DB 60 S-----AGLAGMGGIONEKETMQSLNDRSLASVLDNRSLRLETKIREHLEKKG 111
QY 120 GPQ-VRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSV 178
DB 112 GPQGVDRDWHYFKIIEDLRAQILANSVDNARIVLQIDNARLAADDPRVKYETELAMRQSV 171
QY 179 ENDIHGLRKVIDDNTNITRLQLETEIEALKEELLFMKKNHEEVKGLQAIASSGLTVEVD 238
DB 172 ESDIHGLRKVIDDNTNITRLQLETEIEALKEELLFMKKNHEEVQGLEAASSGLTVEVD 231
QY 239 APKSQDLAKIMADIRAQYDELRARKNREELDKYWSQQIEESTVTTVTTQSAEVAETTLTE 298
DB 232 APKSQDLAKIMADIRAQYDELRARKNREELDKYWSQQIEESTVTTVTTKSAEIRDAETTLTE 291
QY 299 LRRTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQ 358
DB 292 LRRTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQ 351
QY 359 QAQYEALLNIKVKLEAEIATYRRLLEDGEDFNGLDSDNSMOTICTTTRRIVDGVK 418
DB 352 QAQYEALLNIKVKLEAEIATYRRLLEDGEDFNGLDSDNSMOTICTTTRRIVDGVK 411
QY 419 VSETNDTRVLRH 430
DB 412 VSETNDTRVLRH 423

RESULT 15
ABG14106
ID ABG14106 standard; protein; 675 AA.
XX
AC ABG14106;

This Page Blank (uspto)

QY 361 QYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNMOTIQKTTTTRIVDGVKVS 420
Db 361 QYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNMOTIQKTTTTRIVDGVKVS 420
QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430
RESULT 2
US-09-919-497-78
; Sequence 78, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 78
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-78

Query Match 52.4%; Score 1110; DB 4; Length 232;
Best Local Similarity 98.3%; Pred. No. 5.1e-78;
Matches 228; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 199 LETEIEALKEELLFMKKNHEEVKGLQAIASSGLTVEVDAPKSDLAKIMADIRAQYDE 258
Db 1 LETEIEALKEELLFMKKNHEEVKGLQAIASSGLTVEVDAPKSDLAKIMADIRAQYDE 60
QY 259 LARKNREELDKYWSQOIEESTVVTTSQAEVGAETTLTELRTVQSLEIDLDSMRNLKA 318
Db 61 LARKNREELDKYWSQOIEESTVVTTSQAEVGAETTLTELRTVQSLEIDLDSMRNLKA 120
QY 319 SLENSLREVEARYALQMEQINGILLHLESELAQTRAEQORQAQVEYALLNKKVLEAEIA 378
Db 121 SLENSLREVEARYALQMEQINGILLHLESELAQTRAEQORQAQVEYALLNKKVLEAEIA 180
QY 379 TYRRLLEDGEDFNLDSSNMOTIQKTTTTRIVDGVKVSNDTKVLRH 430
Db 181 TYRRLLEDGEDFNLDSSNMOTIQKTTTTRIVDGVKVSNDTKVLRH 232

RESULT 3
US-09-538-092-837
; Sequence 837, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormat Version 0.9
; SEQ ID NO 837
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P02533
US-09-538-092-837

Query Match 43.4%; Score 920.5; DB 4; Length 471;
Best Local Similarity 47.7%; Pred. No. 5.7e-63;
Matches 208; Conservative 80; Mismatches 121; Indels 27; Gaps 9;
QY 17 GSVQAPSY--GARPVS-----AASVYAGAGSGSRISVSSTSFRCGMSG---GLA 64
Db 37 GSCRAPSYGGGLSVSSSRFSSGGAYGLGGYGGGFFSSSSFGSGFGGYYGGGLGAGLG 96
QY 65 TGIAGGLAGMG--ICNEKETMQSINDRLASYLDVRSLSTENRRLESKIREHLEKKGK- 121
Db 97 GFGGGFAGDGLLVGSEKVTMQLNDRLASYLDKVRALAEANADLEVKIRDWYQORPA 156
QY 122 QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVEND 181
Db 157 EIKDYSPEFKTIEDLRNKILTATVDNANVLLQIDNARLAADDFTKYTELNLRMSVEAD 216
QY 182 IHGLRKVIDDNTNITRLQLETEIEALKEELLFMKKNHEEVKGLQAIASSGLTVEVDAPK 241
Db 217 INGLRRVLDLTLARADLEMQIESLKEBELAYLKKNHEEMNALRGQVGGD-VNVEMDAAP 275
QY 242 SQDLAKIMADIRAQYDELDLARKNREELDKYWSQOIEESTVVTTSQAEVGAETTLTELRR 301
Db 276 GVDLSRLNEMRDQYKVAEKNRKDAEWFTEELNREVNATNSELVQSGKSEISELRR 335
QY 302 TVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQINGILLHLESELAQTRAEQORQA 361
Db 336 TMQNLEHLSQSLKASLENSLREETKGRYCNQLAQIQEMTGSVEEQLAQLRCMEQONQ 395
QY 362 EYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNMOTIQKTTTTR-----I 413
Db 396 EYKILLDVKTRLEQEIATYRRLLE-GEDAHLSQSFSSQSSSRDVTSSRQIRTKYMDV 454
QY 414 VDGKVSSETNDTKVLR 429
Db 455 HDGKVS--THEQVLR 468

RESULT 4
US-09-919-172-31
; Sequence 31, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 31
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673545 1709387CDI
US-09-919-172-31

Query Match 41.8%; Score 887; DB 4; Length 456;
Best Local Similarity 44.2%; Pred. No. 2.1e-60;
Matches 199; Conservative 85; Mismatches 126; Indels 40; Gaps 10;
QY 4 TTRSTF---STNYRSI-----GSVQAPSYGARPVSSAASVYAGAGSGSRISVSRST 52
Db 8 TSSSTFGGSTRGGSLLAGGGFGGSLSGGSGSRISASARFVSSGGGGYGGGMVVC 67

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-705-660-18

Query Match 40.8%; Score 865.5; DB 2; Length 432;
Best Local Similarity 44.6%; Pred. No. 8.9e-59;
Matches 189; Conservative 83; Mismatches 115; Indels 37; Gaps 8;

QY 27 RPVSSAASV--YAGAGSGSRISVSRSSTFRGGMGGGLATGIAGGL----- 71
Db 6 RQFTSSSSIKSGSGGLGGSSRTSRLS-----GGLGAGSCLGSGAGGLSTLGGSSSYSCY 61
QY 72 -----AGMGGIQ-----NEKETMOSLNDRLASVLDVRSLETENRRLESKIREHLE 117
Db 62 SFGSGGGYSGSGFGVDGLLAGGEKATQNLNDRLASVLDKVRALAEANTELEVKIRDWYQ 121
QY 118 KKGP-QVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDERVKYETELAMRQ 176
Db 122 RQAPGPARDYSQYRTIEELQNKILTATVDNANILLQIDNARLAADDERVKYETELAMRQ 181
QY 177 SVENDIHGLRKVIDDITNITRLQLETEIEALKEELLFMKKNHEEVEVKGLQAIASSGLTVE 236
Db 182 SVEADINGLRVLDLTLARADLEMQIENLKEELAYLKKHHEEMNALRGQVGE-INVE 240
QY 237 VDAPKSDAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTTSAEVGAETTL 296
Db 241 MDAAPGVDSRLNEMRDQYKMAEKNRKDAEDWFFSKTEELNREVATNSLVQSGKSEI 300
QY 297 TELRRTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAE 356
Db 301 SELRRTMQALEIELOSLMKASLEGNLAETENRYCVQLSQIQGLIGSVVEQLAQLRCM 360
QY 357 ORQAEYEALLNIKVKLEAEIATYRRLLEDGEDENLGDALDSSNSMQTIQTTTRIVDG 416
Db 361 EQQNEQYKILLDVKTRLEQEIATYRRLLE-GEDAHLTQYKKEPVTTROV-RTIVEEVQDG 418
QY 417 KQVS 420
Db 419 KVIS 422

RESULT 7

US-08-989-045-18
Sequence 18, Application US/08989045
Patent No. 6027905
GENERAL INFORMATION:
APPLICANT: KEESEE, SUSAN
APPLICANT: OBAR, ROBERT
APPLICANT: WU, YING-JYE
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
DETECTION OF CERVICAL CANCER
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,045
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-989-045-18

Query Match 40.8%; Score 865.5; DB 3; Length 432;
Best Local Similarity 44.6%; Pred. No. 8.9e-59;
Matches 189; Conservative 83; Mismatches 115; Indels 37; Gaps 8;

QY 27 RPVSSAASV--YAGAGSGSRISVSRSSTFRGGMGGGLATGIAGGL----- 71
Db 6 RQFTSSSSIKSGSGGLGGSSRTSRLS-----GGLGAGSCLGSGAGGLSTLGGSSSYSCY 61
QY 72 -----AGMGGIQ-----NEKETMOSLNDRLASVLDVRSLETENRRLESKIREHLE 117
Db 62 SFGSGGGYSGSGFGVDGLLAGGEKATQNLNDRLASVLDKVRALAEANTELEVKIRDWYQ 121
QY 118 KKGP-QVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDERVKYETELAMRQ 176
Db 122 RQAPGPARDYSQYRTIEELQNKILTATVDNANILLQIDNARLAADDERVKYETELAMRQ 181
QY 177 SVENDIHGLRKVIDDITNITRLQLETEIEALKEELLFMKKNHEEVEVKGLQAIASSGLTVE 236
Db 182 SVEADINGLRVLDLTLARADLEMQIENLKEELAYLKKHHEEMNALRGQVGE-INVE 240
QY 237 VDAPKSDAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTTSAEVGAETTL 296
Db 241 MDAAPGVDSRLNEMRDQYKMAEKNRKDAEDWFFSKTEELNREVATNSLVQSGKSEI 300
QY 297 TELRRTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAE 356
Db 301 SELRRTMQALEIELOSLMKASLEGNLAETENRYCVQLSQIQGLIGSVVEQLAQLRCM 360
QY 357 ORQAEYEALLNIKVKLEAEIATYRRLLEDGEDENLGDALDSSNSMQTIQTTTRIVDG 416
Db 361 EQQNEQYKILLDVKTRLEQEIATYRRLLE-GEDAHLTQYKKEPVTTROV-RTIVEEVQDG 418
QY 417 KQVS 420
Db 419 KVIS 422

RESULT 8

US-09-919-172-9
Sequence 9, Application US/09919172
Patent No. 6673545
GENERAL INFORMATION:
APPLICANT: Paris, Mary
APPLICANT: Turner, Christopher M.
TITLE OF INVENTION: PROSTATE CANCER MARKERS
FILE REFERENCE: PA-0036 US
CURRENT APPLICATION NUMBER: US/09/919,172
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/222,469
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6673545 1454852CD1
US-09-919-172-9

Query Match

40.8%; Score 865.5; DB 4; Length 432;

QY 319 SLENSLREVEARYALQMEQNGILLHLESELAQTRAEQORQAEYEAALLNIKVKLEAIA 378
 Db 306 ALESTLAETARYSSQLAQMCMTITVBAQLAEIRADLERQNEQVQLLDVRARLECEIN 365
 QY 379 TYRRLLEDGEDENL 392
 Db 366 TYRGLLE-SEDSKL 378

RESULT 11
 US-09-067-351-3
 ; Sequence 3, Application US/09067351
 ; Patent No. 5994081
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Baughn, Mariah
 ; TITLE OF INVENTION: HUMAN KERATINS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/067,351
 ; FILING DATE: Herewith
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CERRONE, MICHAEL C.
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0511 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 422 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: BRSTNOT07
 ; CLONE: 2124178
 US-09-067-351-3

Query Match 32.1%; Score 679.5; DB 2; Length 422;
 Best Local Similarity 38.8%; Pred. No. 2e-44;
 Matches 164; Conservative 91; Mismatches 125; Indels 43; Gaps 14;

QY 20 QAPS---YGA-----RPVS--SAASVYAGAGSGSRISVS---RSTSPRGMGSGGLATG 66
 Db 9 QTPSASFHAGGGWGRPRSPFPRAPTVHGAG--GARISLFTTRSCPPPGGSGWGSGRSP 66
 QY 67 IAGGLAGMGGIQNEKETWQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKGK-QVRD 125
 Db 67 LLGG-----NGKATQNLNDRLASYLEKVRALAEANMKLESRLKWHQQRDPGSKKD 118
 QY 126 WSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMQSVENDHGL 185
 Db 119 YSQYEENITHLQEQIVDGKMTNAQIILLIDNARVAVDNFKYENHSFKKDLIEVEGL 178
 QY 186 RKVIDDNTIRLQLETEIEALKELLFMKKNHSEVKGGLQAOIASS-GLTVEVDAPKSOD 244

Db 179 RRTLDNLTIVTTDLEQVEGMRKELILMKKHHEQEME--KHHVPSDFNVNVKVDTPRED 236
 QY 245 LAKIMADIRAQYDELARKNREELDKYWSQ-----IEESTTVVTTQSAEVGAATTLTEL 300
 Db 237 LIKVLDMRQYELIHKHRLDLDTWYKEQSAAMSQEAASPATVQS-----RQGDHLEK 291
 QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEQORQA 360
 Db 292 RTFQALEIDLQTYSTKSALENMLSETQSRYSCKLQDMQEIISHYEEELTQLRHELERQN 351
 QY 361 QEYEALNIKVKLEARIATYRRLLEDGEDENLGDALDSSNSMQTIQ---KTTTRRIVDG 416
 Db 352 NEYQVLLGIKTHLEKITTYRRLLE-GE--SEGTREESKSSMKVSATPKIKAITQETING 408
 QY 417 KWV 419
 Db 409 RLV 411

RESULT 12
 US-09-360-490-3
 ; Sequence 3, Application US/09360490
 ; Patent No. 6221843
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Baughn, Mariah
 ; TITLE OF INVENTION: HUMAN KERATINS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/360,490
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/067,351
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CERRONE, MICHAEL C.
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0511 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 422 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: BRSTNOT07
 ; CLONE: 2124178
 US-09-360-490-3

Query Match 32.1%; Score 679.5; DB 3; Length 422;
 Best Local Similarity 38.8%; Pred. No. 2e-44;
 Matches 164; Conservative 91; Mismatches 125; Indels 43; Gaps 14;

QY 20 QAPS---YGA-----RPVS--SAASVYAGAGSGSRISVS---RSTSPRGMGSGGLATG 66


```

QY      389 -----DFNLGDA-LDSSNSMQTIQKTT-----TRRIVDGKVSETN 423
          ||| : ||| : | | | | | | | | | | | | | | | | | | | |
DB      414 LPNFSSLNLRETNLDSLPLVDTHSKRTFLIKTVETRDGOVINETS 458

Search completed: December 14, 2004, 09:21:55
Job time : 86.4978 secs
```

Qy	348	ELAQTRAEGQRQAQEYVEALNINIKVLEAEIATYRRLLE-----	385
Dd	359	ALQRAKODMARQLREYQELMNVLALDIEIATYRKLEGEBSRLESQMNMNSIHKTTSG	418
Qy	386	-----DGEDFNLDALDSSNSMQTIQKTTTRIV-----	423
Dd	419	YAGGLSSAYGGLTSPCLSYSLGSSFGSGAGSSSFRTSSRAVVVKIEFRDCKLVSESS	478
Qy	424	D 424	
Dd	479	D 479	

```

RESULT 15
US-09-538-092-877
; Sequence 877, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSegformatter Version 0.9
; SEQ ID NO 877
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P08670
US-09-538-092-877

```

Query Match 26.9%; Score 570; DB 4; Length 465;
Best Local Similarity 32.5%; Pred. No. 6.3e-36;
Matches 151; Conservative 96; Mismatches 166; Indels 52; Gaps 13;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 09:07:06 ; Search time 279.782 Seconds
(without alignments)
548.952 Million cell updates/sec

Title: US-10-026-001-1

Perfect score: 2120

Sequence: 1 MSFTTRSTFTSTNYRSLGVSQ.....RRIVDGKVVSETNDTKVLRH 430

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2120	100.0	430	10	US-09-919-039-105
2	2120	100.0	430	10	US-09-913-432-36
3	2120	100.0	430	13	US-10-026-001-1
4	2120	100.0	430	14	US-10-060-036-155
5	2120	100.0	430	14	US-10-174-364-36
6	2120	100.0	430	15	US-10-080-334-265
7	2120	100.0	430	15	US-10-246-583-36
8	2120	100.0	430	16	US-10-689-832-36
9	2120	100.0	430	16	US-10-734-564-117
10	2120	100.0	452	14	US-10-106-698-5102
11	2115	99.8	430	15	US-10-080-334-267
12	2090	98.6	424	15	US-10-080-334-266
13	1835	86.6	375	15	US-10-080-334-268

14	1823.5	86.0	423	15	US-10-080-334-269	Sequence 269, Appl
15	1708.5	80.6	427	10	US-09-813-432-8	Sequence 8, Appli
16	1708.5	80.6	427	14	US-10-174-364-8	Sequence 8, Appli
17	1708.5	80.6	427	15	US-10-246-583-8	Sequence 8, Appli
18	1708.5	80.6	427	16	US-10-689-832-8	Sequence 8, Appli
19	1469	69.3	402	15	US-10-080-334-88	Sequence 88, Appl
20	1362	64.2	627	16	US-10-408-765A-2767	Sequence 2767, Ap
21	1110	52.4	232	9	US-09-919-497-78	Sequence 78, Appl
22	1040.5	49.1	443	16	US-10-408-765A-2578	Sequence 2578, Ap
23	974.5	46.0	231	15	US-10-264-049-3291	Sequence 3291, Ap
24	887	41.8	456	9	US-09-919-172-31	Sequence 31, Appl
25	887	41.8	456	15	US-10-363-616-310	Sequence 310, App
26	875.5	41.3	400	9	US-09-922-217-1115	Sequence 1115, Ap
27	875.5	41.3	400	13	US-10-025-380-1115	Sequence 1115, Ap
28	875.5	41.3	400	17	US-10-733-869A-33	Sequence 33, Appl
29	870.5	41.1	400	16	US-10-734-564-118	Sequence 118, App
30	868	40.9	459	9	US-09-925-298-469	Sequence 469, App
31	868	40.9	459	14	US-10-102-806-469	Sequence 469, App
32	865.5	40.8	432	9	US-09-919-172-9	Sequence 9, Appli
33	865.5	40.8	432	10	US-09-315-355-18	Sequence 18, Appl
34	849	40.0	458	14	US-10-205-823-216	Sequence 216, App
35	841.5	39.7	618	9	US-09-925-300-1381	Sequence 1381, Ap
36	839.5	39.6	593	15	US-10-435-696-50	Sequence 50, Appl
37	821	38.7	494	15	US-10-435-696-51	Sequence 51, Appl
38	815	38.4	424	9	US-09-981-353-52	Sequence 52, Appl
39	815	38.4	424	16	US-10-734-564-119	Sequence 119, App
40	811	38.3	483	14	US-10-316-253-8	Sequence 8, Appli
41	766	36.1	259	15	US-10-108-260A-4615	Sequence 4615, Ap
42	755.5	35.6	312	10	US-09-976-782-32	Sequence 32, Appl
43	755.5	35.6	312	15	US-10-080-334-167	Sequence 167, App
44	755.5	35.6	312	15	US-10-080-334-188	Sequence 188, App
45	755.5	35.6	312	15	US-10-080-334-270	Sequence 270, App

ALIGNMENTS

RESULT 1

US-09-919-039-105
; Sequence 105, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 105
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1272969CD1
US-09-919-039-105

Query Match	100.0%	Score 2120;	DB 10;	Length 430;
Best Local Similarity	100.0%;	Pred. No. 4.1e-132;		
Matches 430;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSFTTRSTFTSTNYRSLGVSQAPSYGARPVSSAASVYAGAGSGSRISVSSTFRGGMCS	60	
Db	1	MSFTTRSTFTSTNYRSLGVSQAPSYGARPVSSAASVYAGAGSGSRISVSSTFRGGMCS	60	
QY	61	GGLATGIAGGLAGGCGIONEKETMQSLNDRSLASYLDRVRSLETENRRLEKKG	120	
Db	61	GGLATGIAGGLAGGCGIONEKETMQSLNDRSLASYLDRVRSLETENRRLEKKG	120	
QY	121	PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN	180	

Db 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKKEELLFMKNHHEEVKGLQQAASSGLTVEVDAP 240
Db 181 DIHGLRKVIDDNTNITRLQLETEIEALKKEELLFMKNHHEEVKGLQQAASSGLTVEVDAP 240
QY 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEVAETTLTEL 300
Db 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEVAETTLTEL 300
QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGQROA 360
Db 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGQROA 360
QY 361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRRIVDGKVV 420
Db 361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRRIVDGKVV 420
QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430

RESULT 2
US-09-813-432-36
; Sequence 36, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: MaJmuder, Kamud
; APPLICANT: Spaderna, Steven K
; APPLICANT: Smithson, Glenda
; APPLICANT: Mezes, Peter S
; APPLICANT: Vernet, Corine A. M.
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729
; CURRENT APPLICATION NUMBER: US/09/813,432
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-432-36

Query Match 100.0%; Score 2120; DB 10; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.le-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTFTSTNYRSLGSGVQAPSGARPVSSAAVYAGAGGSGSRISVSRSSTSTFRGGMGS 60
Db 1 MSFTTRSTFTSTNYRSLGSGVQAPSGARPVSSAAVYAGAGGSGSRISVSRSSTSTFRGGMGS 60
QY 61 GGLATGIAGGLAGMGGIQQNEKETWQSLNDRSLASVLDVRSLRLESENRRLESKIREHLEKKG 120
Db 61 GGLATGIAGGLAGMGGIQQNEKETWQSLNDRSLASVLDVRSLRLESENRRLESKIREHLEKKG 120
QY 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
Db 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKKEELLFMKNHHEEVKGLQQAASSGLTVEVDAP 240
Db 181 DIHGLRKVIDDNTNITRLQLETEIEALKKEELLFMKNHHEEVKGLQQAASSGLTVEVDAP 240
QY 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEVAETTLTEL 300
Db 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEVAETTLTEL 300
QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGQROA 360
Db 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGQROA 360
QY 361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRRIVDGKVV 420
Db 361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNLDSSNSMQTIQKTTTTRRIVDGKVV 420
QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430

RESULT 3
US-10-026-001-1
; Sequence 1, Application US/10026001
; Publication No. US20020122791A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: ANTIGENIC CK-18 COMPOUNDS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS AND METHODS FOR USING SAME
; FILE REFERENCE: GZ 210800
; CURRENT APPLICATION NUMBER: US/10/026,001
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,820
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-001-1

Query Match 100.0%; Score 2120; DB 13; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.le-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTFTSTNYRSLGSGVQAPSGARPVSSAAVYAGAGGSGSRISVSRSSTSTFRGGMGS 60
Db 1 MSFTTRSTFTSTNYRSLGSGVQAPSGARPVSSAAVYAGAGGSGSRISVSRSSTSTFRGGMGS 60
QY 61 GGLATGIAGGLAGMGGIQQNEKETWQSLNDRSLASVLDVRSLRLESENRRLESKIREHLEKKG 120
Db 61 GGLATGIAGGLAGMGGIQQNEKETWQSLNDRSLASVLDVRSLRLESENRRLESKIREHLEKKG 120
QY 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
Db 121 PQVRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
QY 181 DIHGLRKVIDDNTNITRLQLETEIEALKKEELLFMKNHHEEVKGLQQAASSGLTVEVDAP 240
Db 181 DIHGLRKVIDDNTNITRLQLETEIEALKKEELLFMKNHHEEVKGLQQAASSGLTVEVDAP 240

QY 241 KSQDLAKIMADIRAOYDELAARKNEELDKYWSQOIEESTVVTTSQAEVGAETTLTEL 300
Db |||||
QY 241 KSQDLAKIMADIRAOYDELAARKNEELDKYWSQOIEESTVVTTSQAEVGAETTLTEL 300
Db |||||
QY 301 RTVQSLDIDLSMRNLKASLENSLREVEARVALQMEQLNGILLHLESELAQTRAEGORQA 360
Db |||||
QY 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLGDALDSSNSMOTIQKTTTTRIVDGKVS 420
Db |||||
QY 421 ETNDTKVLRH 430
Db |||||

RESULT 4
US-10-060-036-155
; Sequence 155, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 155
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-155

Query Match 100.0%; Score 2120; DB 14; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.1e-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFTTRSTFTNYSLSGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTSFRCGMS 60
Db |||||
QY 1 MSFTTRSTFTNYSLSGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTSFRCGMS 60
Db |||||
QY 61 GGLATGIAGGLAGGGIONEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKG 120
Db |||||
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIQIDNARLAADDPRVKYETELAMRQSVEN 180
Db |||||
QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVKGLQAIASSGLTVEVDAP 240
Db |||||
QY 241 KSQDLAKIMADIRAOYDELAARKNEELDKYWSQOIEESTVVTTSQAEVGAETTLTEL 300
Db |||||
QY 301 RTVQSLDIDLSMRNLKASLENSLREVEARVALQMEQLNGILLHLESELAQTRAEGORQA 360
Db |||||
QY 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLGDALDSSNSMOTIQKTTTTRIVDGKVS 420
Db |||||

QY 421 ETNDTKVLRH 430
Db |||||
QY 421 ETNDTKVLRH 430
Db |||||
RESULT 5
US-10-174-364-36
; Sequence 36, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2
; CURRENT APPLICATION NUMBER: US/10/174,364
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-364-36

Query Match 100.0%; Score 2120; DB 14; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.1e-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFTTRSTFTNYSLSGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTSFRCGMS 60
Db |||||
QY 1 MSFTTRSTFTNYSLSGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTSFRCGMS 60
Db |||||
QY 61 GGLATGIAGGLAGGGIONEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKG 120
Db |||||
QY 61 GGLATGIAGGLAGGGIONEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKG 120
Db |||||
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIQIDNARLAADDPRVKYETELAMRQSVEN 180
Db |||||
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIQIDNARLAADDPRVKYETELAMRQSVEN 180
Db |||||
QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVKGLQAIASSGLTVEVDAP 240
Db |||||
QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEVKGLQAIASSGLTVEVDAP 240
Db |||||
QY 241 KSQDLAKIMADIRAOYDELAARKNEELDKYWSQOIEESTVVTTSQAEVGAETTLTEL 300
Db |||||
QY 241 KSQDLAKIMADIRAOYDELAARKNEELDKYWSQOIEESTVVTTSQAEVGAETTLTEL 300
Db |||||
QY 301 RTVQSLDIDLSMRNLKASLENSLREVEARVALQMEQLNGILLHLESELAQTRAEGORQA 360
Db |||||
QY 301 RTVQSLDIDLSMRNLKASLENSLREVEARVALQMEQLNGILLHLESELAQTRAEGORQA 360
Db |||||

QY 361 QEYEAALLNIKVKLEAEIATYRLLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 420
Db 361 QEYEAALLNIKVKLEAEIATYRLLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 420
QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430

RESULT 6
US-10-080-334-265
; Sequence 265, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shimkets, Richard A
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 265
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-265
Query Match 100.0%; Score 2120; DB 15; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.1e-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFTTRSTFNYSLSGVSQAPSYGARPVSSAAASVYAGAGSGSRISVSRSSTSFRRGMS 60
Db 1 MSFTTRSTFNYSLSGVSQAPSYGARPVSSAAASVYAGAGSGSRISVSRSSTSFRRGMS 60
QY 61 GGLATGIAGLAGMGIGNEKETQSLNDRSLASVLDVRSLRLENNRRLSKIREHLEKKG 120
Db 61 GGLATGIAGLAGMGIGNEKETQSLNDRSLASVLDVRSLRLENNRRLSKIREHLEKKG 120
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
Db 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELFMKNHHEEVKGLQAIASSGLTVEVDAP 240
Db 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELFMKNHHEEVKGLQAIASSGLTVEVDAP 240
QY 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQIIEESTVVTQSAEVAAGAAETTLTEL 300
Db 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQIIEESTVVTQSAEVAAGAAETTLTEL 300
QY 301 RTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGORQA 360
Db 301 RTVQSLEIDLSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGORQA 360
QY 361 QEYEAALLNIKVKLEAEIATYRLLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 420
Db 361 QEYEAALLNIKVKLEAEIATYRLLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 420
QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430

RESULT 7
US-10-246-583-36
; Sequence 36, Application US/10246583
; Publication No. US20040058862A1
; GENERAL INFORMATION:
; APPLICANT: Majumder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28

```
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-246-583-36

Query Match      100.0%; Score 2120; DB 15; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.1e-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTFRGGMGS 60
Db      1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTFRGGMGS 60

QY      61 GGLATGIAGGLAGMGGIQNEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKG 120
Db      61 GGLATGIAGGLAGMGGIQNEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKG 120

QY      121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
Db      121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180

QY      181 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLQAIASSGLTVEVDAP 240
Db      181 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLQAIASSGLTVEVDAP 240

QY      241 KSQDLAKIMADIRAOYDELRKKNREELDKYWSQIIEESTVVTTSQAEVGAETTLTEL 300
Db      241 KSQDLAKIMADIRAOYDELRKKNREELDKYWSQIIEESTVVTTSQAEVGAETTLTEL 300

QY      301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
Db      301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360

QY      361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRRIVDGKVS 420
Db      361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRRIVDGKVS 420

QY      421 ETNDTKVLRH 430
Db      421 ETNDTKVLRH 430
```

```
RESULT 8
US-10-689-832-36
; Sequence 36, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majmuder, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729DIV1
; CURRENT APPLICATION NUMBER: US/10/689,832
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
```

```
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-832-36

Query Match      100.0%; Score 2120; DB 16; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.1e-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTFRGGMGS 60
Db      1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTFRGGMGS 60

QY      61 GGLATGIAGGLAGMGGIQNEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKG 120
Db      61 GGLATGIAGGLAGMGGIQNEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKG 120

QY      121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
Db      121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180

QY      181 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLQAIASSGLTVEVDAP 240
Db      181 DIHGLRKVIDDNTNITRLQLETEIEALKBEELLFMKKNHEEEVKGLQAIASSGLTVEVDAP 240

QY      241 KSQDLAKIMADIRAOYDELRKKNREELDKYWSQIIEESTVVTTSQAEVGAETTLTEL 300
Db      241 KSQDLAKIMADIRAOYDELRKKNREELDKYWSQIIEESTVVTTSQAEVGAETTLTEL 300

QY      301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360
Db      301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGRQA 360

QY      361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRRIVDGKVS 420
Db      361 QEYEALLNIKVKLEAEIATYRRLLEDGEDFNGLDSSNSMQTIQKTTTTRRIVDGKVS 420

QY      421 ETNDTKVLRH 430
Db      421 ETNDTKVLRH 430
```

```
RESULT 9
US-10-734-564-117
; Sequence 117, Application US/10734564
; Publication No. US20040157278A1
; GENERAL INFORMATION:
; APPLICANT: Christopher C Burgess et al
; TITLE OF INVENTION: Detection Methods Using TIMP1
; FILE REFERENCE: 1657/2012
; CURRENT APPLICATION NUMBER: US/10/734,564
; CURRENT FILING DATE: 2003-12-12
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-734-564-117

Query Match      100.0%; Score 2120; DB 16; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.1e-132;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MSFTTRSTFSTNYRSLGVSQAPSYGAPVPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGS 60
 DB 1 MSFTTRSTFSTNYRSLGVSQAPSYGAPVPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGS 60
 QY 61 GGLATGIAGGLAGMGGIQNEKETMQSLNDRSLASVLDVRSLRLENNRRLESKIREHLEKKG 120
 DB 61 GGLATGIAGGLAGMGGIQNEKETMQSLNDRSLASVLDVRSLRLENNRRLESKIREHLEKKG 120
 QY 121 PQVRDWSHYFKIILEDLRAQIFANTVDNARIVLQIDNARLAADFRVKYETELAMRQSVEN 180
 DB 121 PQVRDWSHYFKIILEDLRAQIFANTVDNARIVLQIDNARLAADFRVKYETELAMRQSVEN 180
 QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLQQAIASSGLTVEVDAP 240
 DB 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLQQAIASSGLTVEVDAP 240
 QY 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQIIEESTTQSAEVAETTLTEL 300
 DB 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQIIEESTTQSAEVAETTLTEL 300
 QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAEGQQA 360
 DB 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAEGQQA 360
 QY 361 QEYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 420
 DB 361 QEYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 420
 QY 421 ETNDTKVLRH 430
 DB 421 ETNDTKVLRH 430

RESULT 10

US-10-106-698-5102
 ; Sequence 5102, Application US/10106698
 ; Publication No. US20030109690A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
 ; FILE REFERENCE: PA005PI
 ; CURRENT APPLICATION NUMBER: US/10/106,698
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/157,137
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: US 60/163,280
 ; PRIOR FILING DATE: 1999-11-03
 ; NUMBER OF SEQ ID NOS: 8564
 ; SOFTWARE: Patent In Ver. 3.0
 ; SEQ ID NO 5102
 ; LENGTH: 452
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-106-698-5102

Query Match 100.0%; Score 2120; DB 14; Length 452;
 Best Local Similarity 100.0%; Pred. No. 4.3e-132;
 Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSFTTRSTFSTNYRSLGVSQAPSYGAPVPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGS 60
 DB 23 MSFTTRSTFSTNYRSLGVSQAPSYGAPVPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGS 82
 QY 61 GGLATGIAGGLAGMGGIQNEKETMQSLNDRSLASVLDVRSLRLENNRRLESKIREHLEKKG 120
 DB 83 GGLATGIAGGLAGMGGIQNEKETMQSLNDRSLASVLDVRSLRLENNRRLESKIREHLEKKG 142
 QY 121 PQVRDWSHYFKIILEDLRAQIFANTVDNARIVLQIDNARLAADFRVKYETELAMRQSVEN 180
 DB 143 PQVRDWSHYFKIILEDLRAQIFANTVDNARIVLQIDNARLAADFRVKYETELAMRQSVEN 202

QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLQQAIASSGLTVEVDAP 240
 DB 203 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKKNHEEEVKGLQQAIASSGLTVEVDAP 262
 QY 241 KSQDLAKIMADIRAQYDELARKNREELDKYWSQIIEESTTQSAEVAETTLTEL 300
 DB 263 KSQDLAKIMADIRAQYDELARKNREELDKYWSQIIEESTTQSAEVAETTLTEL 322
 QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAEGQQA 360
 DB 323 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNLGILLHLESELAQTRAEGQQA 382
 QY 361 QEYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 420
 DB 383 QEYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNSMOTIQKTTTTRIVDGKVS 442
 QY 421 ETNDTKVLRH 430
 DB 443 ETNDTKVLRH 452

RESULT 11

US-10-080-334-267
 ; Sequence 267, Application US/10080334
 ; Publication No. US20040002584A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pena, Carol E. A.
 ; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: Li, Li
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Vernet, Corine A. M.
 ; APPLICANT: Malyankar, Uriel M.
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Gusev, Vladimir Y
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Boldog, Ferenc L
 ; APPLICANT: Furtak, Katarzyna T
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Gangolli, Esha A
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Baumgartner, Jason C.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Spaderna, Steven K
 ; APPLICANT: Zerhusen, Bryan D
 ; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
 ; FILE REFERENCE: 21402-275
 ; CURRENT APPLICATION NUMBER: US/10/080,334
 ; CURRENT FILING DATE: 2002-02-21
 ; PRIOR APPLICATION NUMBER: 60/270,523
 ; PRIOR FILING DATE: 2001-02-21
 ; PRIOR APPLICATION NUMBER: 60/322,712
 ; PRIOR FILING DATE: 2001-09-17
 ; PRIOR APPLICATION NUMBER: 60/311,980
 ; PRIOR FILING DATE: 2001-08-13
 ; PRIOR APPLICATION NUMBER: 60/330,307
 ; PRIOR FILING DATE: 2001-10-18
 ; PRIOR APPLICATION NUMBER: 60/278,796
 ; PRIOR FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 60/281,521
 ; PRIOR FILING DATE: 2001-04-04
 ; PRIOR APPLICATION NUMBER: 60/276,677
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/311,595
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/270,220
 ; PRIOR FILING DATE: 2001-02-21
 ; PRIOR APPLICATION NUMBER: 60/274,295
 ; PRIOR FILING DATE: 2001-03-08

;; PRIOR APPLICATION NUMBER: 60/318,526
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/286,548
;; PRIOR FILING DATE: 2001-04-25
;; PRIOR APPLICATION NUMBER: 60/291,765
;; PRIOR FILING DATE: 2001-05-17
;; PRIOR APPLICATION NUMBER: 60/270,797
;; PRIOR FILING DATE: 2001-02-23
;; PRIOR APPLICATION NUMBER: 60/276,400
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/270,810
;; PRIOR FILING DATE: 2001-02-23
;; NUMBER OF SEQ ID NOS: 388
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 267
;; LENGTH: 430
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-080-334-267

Query Match 99.8%; Score 2115; DB 15; Length 430;
Best Local Similarity 99.8%; Pred. No. 8.7e-132;
Matches 429; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFTTRSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSSTFRGGMG 60
Db 1 MSFTTRSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSSTFRGGMG 60

QY 61 GGLATGIAGLAGMGGIONEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEK 120
Db 61 GGLATGIAGLAGMGGIONEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEK 120

QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180
Db 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180

QY 181 DIHGLRKVIDDTNITRLQLETEIEALKEBLLFMKKNHEEVEVKGLQAIASSGLTVEVDAP 240
Db 181 DIHGLRKVIDDTNITRLQLETEIEALKEBLLFMKKNHEEVEVKGLQAIASSGLTVEVDAP 240

QY 241 KSQDLAKIMADIRAOYDELARKNREEDKYWSQQIEESTTQSAEVEGAAETTLTEL 300
Db 241 KSQDLAKIMADIRAOYDELARKNREEDKYWSQQIEESTTQSAEVEGAAETTLTEL 300

QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQORQA 360
Db 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQORQA 360

QY 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNGLDSSNSMOTIKTTTTRIVDGKVS 420
Db 361 QEYEAALLNIKVKLEAEIATYRRLLEDGEDFNGLDSSNSMOTIKTTTTRIVDGKVS 420

QY 421 ETNDTKVLRH 430
Db 421 ETNDTKVLRH 430

RESULT 12
US-10-080-334-266
; Sequence 266, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shimkets, Richard A
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J

;; APPLICANT: Boldog, Ferenc L
;; APPLICANT: Furtak, Katarzyna
;; APPLICANT: Tchernev, Velizar T
;; APPLICANT: Patturajan, Meera
;; APPLICANT: Gangolli, Esha A
;; APPLICANT: Padigaru, Muralidhara
;; APPLICANT: Liu, Xiaohong
;; APPLICANT: Baumgartner, Jason C.
;; APPLICANT: Gerlach, Valerie
;; APPLICANT: Spaderna, Steven K
;; APPLICANT: Zerhusen, Bryan D
;; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
;; TITLE OF INVENTION: Using the Same
;; FILE REFERENCE: 21402-275
;; CURRENT APPLICATION NUMBER: US/10/080,334
;; CURRENT FILING DATE: 2002-02-21
;; PRIOR APPLICATION NUMBER: 60/270,523
;; PRIOR FILING DATE: 2001-02-21
;; PRIOR APPLICATION NUMBER: 60/322,712
;; PRIOR FILING DATE: 2001-09-17
;; PRIOR APPLICATION NUMBER: 60/311,980
;; PRIOR FILING DATE: 2001-08-13
;; PRIOR APPLICATION NUMBER: 60/330,307
;; PRIOR FILING DATE: 2001-10-18
;; PRIOR APPLICATION NUMBER: 60/278,796
;; PRIOR FILING DATE: 2001-03-26
;; PRIOR APPLICATION NUMBER: 60/281,521
;; PRIOR FILING DATE: 2001-04-04
;; PRIOR APPLICATION NUMBER: 60/276,677
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/311,595
;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: 60/270,220
;; PRIOR FILING DATE: 2001-02-21
;; PRIOR APPLICATION NUMBER: 60/274,295
;; PRIOR FILING DATE: 2001-03-08
;; PRIOR APPLICATION NUMBER: 60/318,526
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/286,548
;; PRIOR FILING DATE: 2001-04-25
;; PRIOR APPLICATION NUMBER: 60/291,765
;; PRIOR FILING DATE: 2001-05-17
;; PRIOR APPLICATION NUMBER: 60/270,797
;; PRIOR FILING DATE: 2001-02-23
;; PRIOR APPLICATION NUMBER: 60/276,400
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/270,810
;; PRIOR FILING DATE: 2001-02-23
;; NUMBER OF SEQ ID NOS: 388
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 266
;; LENGTH: 424
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-080-334-266

Query Match 98.6%; Score 2090; DB 15; Length 424;
Best Local Similarity 100.0%; Pred. No. 3.8e-130;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 STFTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGGLATG 66
Db 1 STFTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGGLATG 60

QY 67 IAGGLAGMGGIONEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKGPQVRDW 126
Db 61 IAGGLAGMGGIONEKETMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKGPQVRDW 120

QY 127 SHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVENDIHGLR 186
Db 121 SHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVENDIHGLR 180

QY 187 KVIDDTNITRLQLETEIEALKEBLLFMKKNHEEVEVKGLQAIASSGLTVEVDAPKSDLA 246

Db 181 KVDDTNRQLQTEIEALKKEELLFMKKHHEEVKGLQAIASSGLTVEVDAPXSQDLA 240
QY 247 KIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEYVGAATITELRRVQSL 306
Db 241 KIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEYVGAATITELRRVQSL 300
QY 307 EIDLDSENLKASLENSLREVEARYALQMEQNGILLHLESLAQTRAEGORQAQYEAL 366
Db 301 EIDLDSENLKASLENSLREVEARYALQMEQNGILLHLESLAQTRAEGORQAQYEAL 360
QY 367 LNIKVKLEAEIATYRRLLEDGEDFNLGDALDSSNSMOTIQKTTTRIVDGVSETNDTK 426
Db 361 LNIKVKLEAEIATYRRLLEDGEDFNLGDALDSSNSMOTIQKTTTRIVDGVSETNDTK 420
QY 427 VLRH 430
Db 421 VLRH 424

RESULT 13

US-10-080-334-268
Sequence 268, Application US/10080334
Publication No. US20040002584A1

GENERAL INFORMATION:

APPLICANT: Shimkets, Richard A.
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A.
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Guo, Xiaojia
APPLICANT: Gusev, Vladimir Y
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Furtak, Katarzyna T
APPLICANT: Tchernev, Velizar
APPLICANT: Patturajan, Meera
APPLICANT: Gangolli, Esha A
APPLICANT: Padigar, Muralidhara
APPLICANT: Liu, Xiaohong
APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zerhusen, Bryan D

TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of

TITLE OF INVENTION: Using the Same

FILE REFERENCE: 21402-275

CURRENT APPLICATION NUMBER: US/10/080,334

CURRENT FILING DATE: 2002-02-21

PRIOR APPLICATION NUMBER: 60/270,523

PRIOR FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: 60/322,712

PRIOR FILING DATE: 2001-09-17

PRIOR APPLICATION NUMBER: 60/311,980

PRIOR FILING DATE: 2001-08-13

PRIOR APPLICATION NUMBER: 60/330,307

PRIOR FILING DATE: 2001-10-18

PRIOR APPLICATION NUMBER: 60/278,796

PRIOR FILING DATE: 2001-03-26

PRIOR APPLICATION NUMBER: 60/281,521

PRIOR FILING DATE: 2001-04-04

PRIOR APPLICATION NUMBER: 60/276,677

PRIOR FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 60/311,595

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: 60/270,220

PRIOR FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: 60/274,295

PRIOR FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: 60/318,526

PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 268
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
US-10-080-334-268

Query Match

86.6%; Score 1835; DB 15; Length 375;

Best Local Similarity 100.0%; Pred. No. 2.4e-113;

Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 GSGGLATGIAGGLAGGIONEKETWQSLNDRLASYLDVRSLETENRRLESKIREHLEK 118
Db 4 GSGGLATGIAGGLAGGIONEKETWQSLNDRLASYLDVRSLETENRRLESKIREHLEK 63
QY 119 KGPOVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMQSV 178
Db 64 KGPOVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMQSV 123
QY 179 ENDIHGLEKVIDDTNITRLQLETEIEALKKEELLFMKKHHEEVKGLQAIASSGLTVEVD 238
Db 124 ENDIHGLEKVIDDTNITRLQLETEIEALKKEELLFMKKHHEEVKGLQAIASSGLTVEVD 183
QY 239 APKSQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEYVGAATITL 298
Db 184 APKSQDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTVVTQSAEYVGAATITL 243
QY 299 LRRTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESLAQTRAEGOR 358
Db 244 LRRTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQNGILLHLESLAQTRAEGOR 303
QY 359 QAQYEALLNIKVKLEABIAIATYRRLLEDGEDFNLGDALDSSNSMOTIQKTTTRIVDGV 418
Db 304 QAQYEALLNIKVKLEABIAIATYRRLLEDGEDFNLGDALDSSNSMOTIQKTTTRIVDGV 363
QY 419 VSETNDTKVLRH 430
Db 364 VSETNDTKVLRH 375

RESULT 14

US-10-080-334-269

Sequence 269, Application US/10080334

Publication No. US20040002584A1

GENERAL INFORMATION:

APPLICANT: Pena, Carol E. A.
APPLICANT: Shimkets, Richard A
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A.
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M
APPLICANT: Guo, Xiaojia
APPLICANT: Gusev, Vladimir Y
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Furtak, Katarzyna T
APPLICANT: Tchernev, Velizar
APPLICANT: Patturajan, Meera
APPLICANT: Gangolli, Esha A

APPLICANT: Padigaru, Muralidhara
APPLICANT: Liu, Xiaohong
APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zethusen, Bryan D
TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
TITLE OF INVENTION: Using the Same
FILE REFERENCE: 21402-275
CURRENT APPLICATION NUMBER: US/10/080,334
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/330,307
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/278,796
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/281,521
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/276,677
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/311,595
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/270,220
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/274,295
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/318,526
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 269
LENGTH: 423
TYPE: PRT
ORGANISM: Mus musculus
US-10-080-334-269
Query Match 86.0%; Score 1823.5; DB 15; Length 423;
Best Local Similarity 86.3%; Pred. No. 1.6e-112;
Matches 373; Conservative 26; Mismatches 22; Indels 11; Gaps 4;
QY 1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSSTFRGGMG 59
Db 1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSV-WGGSVG 59
QY 60 SCGLATGIAGGLAGMGGIQQEKEKTMQSLNDRLASVLDVRSLETENRLESKIREHLEKK 119
Db 60 S-----AGLAGMGGIQQEKEKTMQSLNDRLASVLDVRSLETENRLESKIREHLEKK 111
QY 120 GPQ-VRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSV 178
Db 112 GPQGVDRDWHGVFKIIEDLRAQIFANSVDNARIVLQIDNARLAADDPRVKYETELAMRQSV 171
QY 179 ENDIHGLRKVIDDNTNITRLOLETEIEALKBELEKFMKKNHEEVKGLQAQIASGLTVEVD 238
Db 172 ESDIHGLRKVVDDNTNITRLOLETEIEALKBELEKFMKKNHEEVQGLEAQIASGLTVEVD 231
QY 239 APKSQDLAKIMADIRAQYDEBLARKNREELDKYWSQQIEESTTVVTTKSAEIRDAETTLTE 298.

Db 232 APXSQDLSKIMADIRAQYEAALAQKNREELDKYWSQQIEESTTVVTTKSAEIRDAETTLTE 291
QY 299 LRRTVQSLLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQOR 358
Db 292 LRRTLQTLLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEQOR 351
QY 359 QAQYEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDGALDSSNSMQTIQKTTTRRIVDGKV 418
Db 352 QAQYEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDGALDSSNSMQTIQKTTTRKIVDGRV 411
QY 419 VSETNDTKVLRH 430
Db 412 VSETNDTRVLRH 423
RESULT 15
US-09-813-432-8
Sequence 8, Application US/09813432
Publication No. US20030148485A1
GENERAL INFORMATION:
APPLICANT: Taupier Jr., Raymond J
APPLICANT: Majmuder, Kamud
APPLICANT: Spaderna, Steven K
APPLICANT: Smithson, Glenda
APPLICANT: Mezes, Peter S
APPLICANT: Vernet, Corine A. M.
TITLE OF INVENTION: No. US20030148485A1 Polypeptides and Amino Acids Encoding Same
FILE REFERENCE: 15966-729
CURRENT APPLICATION NUMBER: US/09/813,432
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/193,843
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 427
TYPE: PRT
ORGANISM: Homo sapiens
US-09-813-432-8
Query Match 80.6%; Score 1708.5; DB 10; Length 427;
Best Local Similarity 82.3%; Pred. No. 6.7e-105;
Matches 354; Conservative 30; Mismatches 43; Indels 3; Gaps 3;
QY 1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGS 60
Db 1 MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSSTFRGGMGS 60
QY 61 GGLATGIAGGLAGMGGIQQEKEKTMQSLNDRLASVLDVRSLETENRLESKIREHLEKK 120
Db 61 GGLATEMAGGLAEMGGIQQEKEKTMQSLNDHL-DYLDVRNLETDENWRLESKIQEYLEKR- 118
QY 121 PQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSVEN 180

Db 119 PHVRDWGCHYFKTIKELRAQIFANTVDNVHIILOIDNARLAADDFRVXYETELAMRQSVES 178
QY 181 DIHGLRKVIDDTNITRLQLETEIEALKKEELLFMKNHHEEVKGLQAOIASSGLTVEVDAP 240
Db 179 NIHGLCKVIDDTNVTLLQLETEMGALKKEELLFMKNHHEEVKGLQOQIANSGLAVEVDAP 238
QY 241 KSQDLAKIMADIRAQYDELAARKNREELDKYWSQQIEESTTVVTTQSAEVGAAETTLTEL 300
Db 239 KSQVLAKVMADIRAQYDELSQKNSEKLGKYWSQQTEESTTVVTTTHSAKVRRAEMT-TEL 297
QY 301 RTVQSLEIDLDSMRNLKASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQROA 360
Db 298 RTVQCLEIDLDSMRNLKTSLENSLREVEARYALQMEQLNRILLYLESKLAQNWAEGQKV 357
QY 361 QBYEALLNIVKLEAEIATYRRLLEDGEDFNLGDALDSSNSMQTIQKTTTRIVDGKVV 420
Db 358 QBYKDLLNIRVKLEAEIATYRRLLEDSEGLNLGDALDSSNSMQTIQKTTTRIVDSKVV 417
QY 421 ETNDTKVLRH 430
Db 418 EISDTKVLRH 427

Search completed: December 14, 2004, 09:27:00
Job time : 282.782 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31; Search time 339.869 Seconds
(without alignments)
727.960 Million cell updates/sec

Title: US-10-026-001-1

Perfect score: 2120

Sequence: 1 MSFTTRSTFTSNYRSLGSVQ.....RRIVDGKVVSETNDTKVLRH 430

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Uniprot_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2115	99.8	429	1 K1CR_HUMAN	P05783 homo sapien
2	1812.5	85.5	422	1 K1CR_MOUSE	P05784 mus musculus
3	1285	60.6	432	2 Q7SY65	Q7SY65 xenopus lae
4	1250.5	59.0	428	2 Q6INH6	Q6INH6 xenopus lae
5	1250.5	59.0	428	2 AAH72305	AAH72305 xenopus 1
6	1250.5	59.0	436	2 Q8AVI2	Q8AVI2 xenopus lae
7	1249.5	58.9	429	2 Q6P864	Q6P864 xenopus tro
8	1249.5	58.9	429	2 AAH61366	AAH61366 xenopus t
9	1137.5	53.7	368	1 K1CR_XENLA	P08802 xenopus lae
10	1135.5	53.6	435	2 Q7SYF8	Q7SYF8 acipenser b
11	1093.5	51.6	431	2 Q7ZTS4	Q7ZTS4 brachydanio
12	1092.5	51.5	431	2 Q6P042	Q6P042 brachydanio
13	1092.5	51.5	431	2 AAH65848	AAH65848 brachydan
14	1082.5	51.1	438	2 O57607	O57607 oncorhynch
15	1036.5	48.9	415	2 O57611	O57611 scyllorhinu
16	1032.5	48.7	435	2 Q07427	Q07427 carassius a
17	1015.5	47.9	405	2 Q6PHE4	Q6PHE4 brachydanio
18	1015.5	47.9	405	2 AAH56584	AAH56584 brachydan
19	1013	47.8	359	2 Q7ZT44	Q7ZT44 brachydanio
20	1012.5	47.8	405	2 Q6NVL7	Q6NVL7 brachydanio
21	1012.5	47.8	405	2 AAH66541	AAH66541 brachydan
22	969.5	45.7	438	2 Q6DHU3	Q6DHU3 brachydanio
23	953.5	45.0	458	2 Q6DKC7	Q6DKC7 xenopus lae
24	935.5	44.1	452	2 Q6IFU7	Q6IFU7 rattus norv
25	926.5	43.7	452	2 Q6SEK1	Q6SEK1 mus musculu
26	926.5	43.7	452	2 Q6IFX2	Q6IFX2 mus musculu
27	926.5	43.7	452	2 AAR22526	AAR22526 mus muscu
28	926	43.7	412	2 Q6P7K6	Q6P7K6 xenopus tro
29	926	43.7	412	2 AAH61624	AAH61624 xenopus t
30	924.5	43.6	472	2 AAP35850	AAP35850 homo sapi
31	920.5	43.4	471	1 K1CN_HUMAN	P02533 homo sapien

32 919.5 43.4 467 2 Q6PVZ1 Q6pvz1 gallus gall
33 919.5 43.4 467 2 AAS92202 Aas92202 gallus ga
34 916.5 43.2 484 1 K1CN_MOUSE Q61781 mus musculu
35 915 43.2 412 2 Q6IRP5 Q6irp5 xenopus lae
36 915 43.2 412 2 AAH70682 Aah70682 xenopus 1
37 907.5 42.8 456 2 Q6PEQ0 Q6peg0 mus musculu
38 907.5 42.8 456 2 AAH57934 Aah57934 mus muscu
39 907 42.8 485 2 Q6IFV1 Q6ifv1 rattus norv
40 904.5 42.7 452 1 K1CO_MOUSE Q61414 mus musculu
41 900.5 42.5 438 2 Q90W74 Q9ow74 oncorhynch
42 900 42.5 447 2 Q6IFV3 Q6ifv3 rattus norv
43 891.5 42.1 472 1 K1CP_HUMAN P08779 homo sapien
44 889.5 42.0 399 1 K1CS_BOVIN P08728 bos taurus
45 887 41.8 456 1 K1CO_HUMAN P19012 homo sapien

ALIGNMENTS

RESULT 1

K1CR_HUMAN STANDARD; PRT; 429 AA.
ID K1CR_HUMAN
AC P05783; Q9BW26;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Keratin, type I cytoskeletal 18 (Cytokeratin 18) (K18) (CK 18).
GN Name=KRT18; Synonyms=CYK18;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87134778; PubMed=2434380;
RA Oshima R.G., Millan J.L., Cecena G.;
RT "Comparison of mouse and human Keratin 18: a component of intermediate filaments expressed prior to implantation."
RL Differentiation 33:61-68(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix, Colon, Placenta, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 6-429 FROM N.A.
RX MEDLINE=87134779; PubMed=2434381;
RA Leube R.E., Bosch F.X., Romano V., Zimbelmann R., Hofler H.,
RA Franke W.W.;
RT "Cytokeratin expression in simple epithelia. III. Detection of mRNAs encoding human cytokeratins nos. 8 and 18 in normal and tumor cells by hybridization with cDNA sequences in vitro and in situ."
RL Differentiation 33:69-85(1986).
RN [4]

RP SEQUENCE OF 198-429 FROM N.A.
RX MEDLINE=86193258; PubMed=2422083;
RA Romano V., Hatzfeld M., Magin T.M., Zimbelmann R., Franke W.W.,
RA Maier G., Ponstingl H.;
RT "Cyokeratin expression in simple epithelia. I. Identification of mRNA
RT coding for human cytokeratin no. 18 by a cDNA clone.";
RL Differentiation 30:244-253(1986).
RN [5]
RP SEQUENCE OF 1-166 FROM N.A.
RX MEDLINE=88246424; PubMed=2454392;
RA Kulesh D.A., Oshima R.G.;
RT "Cloning of the human keratin 18 gene and its expression in
RT nonepithelial mouse cells.";
RL Mol. Cell. Biol. 8:1540-1550(1988).
RN [6]
RP PARTIAL SEQUENCE.
RC TISSUE=Colon carcinoma;
RX MEDLINE=97295306; PubMed=9150948;
RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins.";
RL Electrophoresis 18:605-613(1997).
RN [7]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=95263524; PubMed=7538124;
RA Ku N.-O., Omary M.B.;
RT "Identification and mutational analysis of the glycosylation sites of
RT human keratin 18.";
RL J. Biol. Chem. 270:11820-11827(1995).
RN [8]
RP VARIANT CRYPTOGENIC CIRRHOSIS LEU-127.
RX MEDLINE=97148766; PubMed=9011570;
RA Ku N.-O., Wright T.L., Terrault N.A., Gish R., Omary M.B.;
RT "Mutation of human keratin 18 in association with cryptogenic
RT cirrhosis.";
RL J. Clin. Invest. 99:19-23(1997).
CC -!- SUBUNIT: Heterotetramer of two type I and two type II keratins.
CC Keratin 18 associates with keratin 8.
CC -!- DISEASE: Defects in KRT18 are a cause of cryptogenic cirrhosis
CC [MIM:215600].
CC -!- MISCELLANEOUS: There are two types of cytoskeletal and
CC 55 and 56-70 kDa, respectively).
CC -!- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF179904; AA59461.1; -
DR EMBL; BC000180; AAH00180.1; -
DR EMBL; BC000698; AAH00698.1; -
DR EMBL; BC004253; AAH04253.1; -
DR EMBL; BC008636; AAH08636.1; -
DR EMBL; BC020382; AAH20382.1; -
DR EMBL; X12876; CAA31369.1; -
DR EMBL; X12881; CAA31375.1; -
DR PIR; S05481; S05481.
DR HSSP; P08670; 1GK7.
DR IntAct; P05783; -
DR GlycoSuiteDB; P05783; -
DR SWISS-2DPAGE; P05783; HUMAN.
DR PMMA-2DPAGE; P05783; -
DR Sienna-2DPAGE; P05783; -
DR Genew; HGNC:6430; KRT18.
DR MIM; 148070; -
DR MIM; 215600; -
DR GO; GO:0005882; C:intermediate filament; TAS.
DR GO; GO:0009653; P:morphogenesis; TAS.

DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PRO1248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; 1.
KW Acetylation; Coiled coil; Direct protein sequencing; Disease mutation;
KW Glycoprotein; Intermediate filament; Keratin.
FT INIT MET 0 0
FT MOD RES 1 1 N-acetylserine.
FT DOMAIN 1 78 Head.
FT DOMAIN 79 386 Rod.
FT DOMAIN 387 429 Tail.
FT DOMAIN 79 114 Coil 1A.
FT DOMAIN 115 131 Linker 1.
FT DOMAIN 132 223 Coil 1B.
FT DOMAIN 224 247 Linker 12.
FT DOMAIN 248 386 Coil 2.
FT SITE 270 270 Stutter.
FT SITE 330 330 Stutter.
FT CARBOHYD 29 29 O-linked (GlcNAc).
FT CARBOHYD 30 30 /FTID=CAR 000175.
FT CARBOHYD 48 48 /FTID=CAR 000193.
FT CARBOHYD 127 127 O-linked (GlcNAc).
FT VARIANT 127 127 /FTID=CAR 000194.
FT H -> L (in cryptogenic cirrhosis;
FT interferes with the ability to form normal
FT filaments).
FT /FTID=VAR 003852.
FT Y -> H (in Ref. 2; AAH00698).
FT E -> Q (in Ref. 4).
FT A -> S (in Ref. 4).
FT D -> R (in Ref. 4).
FT S -> R (in Ref. 4).
SQ SEQUENCE 429 AA; 47926 MW; 7172B9D6A39933F8 CRC64;

Query Match 99.8%; Score 2115; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 9e-97;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SFTTRSTFTNYSRLSGSVQAPSYGAPVSSAASVYAGAGSGSRISVSRSSTFRGGMSSG 61
Db 1 SFTTRSTFTNYSRLSGSVQAPSYGAPVSSAASVYAGAGSGSRISVSRSSTFRGGMSSG 60

QY 62 GLATGIAGGLAGMGGIQNEKETMOSLNDRLASVLDVRVSRLETENRRLESKIREHLEKGP 121
Db 61 GLATGIAGGLAGMGGIQNEKETMOSLNDRLASVLDVRVSRLETENRRLESKIREHLEKGP 120

QY 122 QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADPRVKYETELAMRQSVEND 181
Db 121 QVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADPRVKYETELAMRQSVEND 180

QY 182 IHGLRKVIDDTNITRLQLETEIEALKEELLEFMKNHHEEVKGLQAIASSGLTVEVDAPK 241
Db 181 IHGLRKVIDDTNITRLQLETEIEALKEELLEFMKNHHEEVKGLQAIASSGLTVEVDAPK 240

QY 242 SQDLAKIMADIRAQYDELARKNREELDKYWSQQIIEESTTVVTTSQSAEVGAETTLTLRLR 301
Db 241 SQDLAKIMADIRAQYDELARKNREELDKYWSQQIIEESTTVVTTSQSAEVGAETTLTLRLR 300

QY 302 TVQSLEIDLDSMRNLKASLENSLRVEARYALQWELQNLGILLHLESELAQTRAEQORQAQ 361
Db 301 TVQSLEIDLDSMRNLKASLENSLRVEARYALQWELQNLGILLHLESELAQTRAEQORQAQ 360

QY 362 EYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNSNQTIQKTTTTRRIDVGKVVSE 421
Db 361 EYEALLNKKVLEAEIATYRRLLEDGEDFNLDSSNSNQTIQKTTTTRRIDVGKVVSE 420

QY 422 TNDTKVLRH 430
Db 421 TNDTKVLRH 429

```
RESULT 2
K1CR_MOUSE STANDARD; PRT; 422 AA.
ID K1CR_MOUSE 061766;
AC P05784; Q61766;
DT 01-NOV-1988 (Rel. 09, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Keratin, type I cytoskeletal 18 (Cytokeratin 18) (Cytokeratin endo B)
DE (Keratin D).
GN Name=Krt18; Synonyms=Krt1-18, Kerd;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89196920; PubMed=2467843;
RA Ichinose Y., Morita T., Zhang F., Srimahasongram S., Tondella M.L.C.,
RA Matsumoto M., Nozaki M., Matsushiro A.;
RT "Nucleotide sequence and structure of the mouse cytoke-
RT gene.";
RL Gene 70:85-95 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86085876; PubMed=2416755;
RA Singer P.A., Trevor K., Oshima R.G.;
RT "Molecular cloning and characterization of the Endo B cytoke-
RT rat osteoblast protein induced during in vitro differentiation of F9
RT teratocarcinoma cells.";
RL J. Biol. Chem. 261:538-547 (1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86085876; PubMed=2416755;
RA Singer P.A., Trevor K., Oshima R.G.;
RT "Molecular cloning and characterization of the Endo B cytoke-
RT rat osteoblast protein induced during in vitro differentiation of F9
RT teratocarcinoma cells.";
RL J. Biol. Chem. 261:538-547 (1986).
RN [4]
RP SEQUENCE OF 1-131 FROM N.A.
RX MEDLINE=88255838; PubMed=2454868;
RA Oshima R.G., Trevor K., Shevinsky L.H., Ryder O.A., Cecena G.;
RT "Identification of the gene coding for the Endo B murine cytoke-
RT rat osteoblast protein induced during in vitro differentiation of F9
RT cells.";
RL Genes Dev. 2:505-516 (1988).
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC -1- Keratin 18 associates with keratin 8.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin, I (acidic) and II (neutral to basic) (40-
CC 55 and 56-70 kDa, respectively).
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
RESULT 3
Q7SY65
ID Q7SY65 PRELIMINARY; PRT; 432 AA.
AC Q7SY65;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Krt18-prov protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]

KW Acetylation; Coiled coil; Glycoprotein; Intermediate filament;
KW Keratin.
FT INIT_MET -0 0 By similarity.
FT MOD_RES 1 1 N-acetylserine (By similarity).
FT DOMAIN 1 70 Head.
FT DOMAIN 71 379 Rod.
FT DOMAIN 380 422 Tail.
FT DOMAIN 71 106 Coil 1A.
FT DOMAIN 107 124 Linker 1.
FT DOMAIN 125 216 Coil 1B.
FT DOMAIN 217 240 Linker 12.
FT DOMAIN 241 379 Coil 2.
FT SITE 263 263 Stutter.
FT SITE 323 323 Stutter.
FT CARBOHYD 30 30 O-linked (GlcNAc) (By similarity).
FT CARBOHYD 31 31 O-linked (GlcNAc) (By similarity).
FT CARBOHYD 49 49 O-linked (GlcNAc) (By similarity).
FT CONFLICT 133 133 L -> F (in Ref. 1).
FT CONFLICT 243 243 D -> N (in Ref. 2).
FT CONFLICT 252 252 A -> G (in Ref. 2).
SQ SEQUENCE 422 AA; 47373 MW; 4D5B0E9C7732F2F3 CRC64;

Query Match 85.5%; Score 1812.5; DB 1; Length 422;
Best Local Similarity 86.1%; Pred. No. 7.6e-82;
Matches 371; Conservative 26; Mismatches 23; Indels 11; Gaps 4;

QY 2 SFTTRS-TFSTNYSLSGVQAPSGYAGPVSAAASVYAGAGSGSRISVSRSTSRFGMGS 60
Db 1 SFTTRSTTFSTNYSLSGVQAPSGYAGPVSAAASVYAGAGSGSRISVSRSV-WGGSVGS 59
QY 61 GGLATGIAGGLAGGGIQTETKQSLNDRLASLDVRSLETENRRLESKIREHLEKKG 120
Db 60 -----AGLAGGGIQTETKQSLNDRLASLDVRSLETENRRLESKIREHLEKKG 111
QY 121 PQ-VRDWSHYFKIIEEDLRAQIFANTVDNARIVLQIDNARLAADDPRVKYETELAMRQSV 179
Db 112 PQGVDRWGHYFKIIEEDLRAQILANSVDNARIVLQIDNARLAADDPRVKYETELAMRQSV 171
QY 180 NDIHGLRKVIDDTNITRLQLETEALKEELLFMKNHEEVKGLQAIASSGLTVEVDA 239
Db 172 SDIHGLRKVVDDTNITRLQLETEALKEELLFMKNHEEVKGLQAIASSGLTVEVDA 231
QY 240 PKSQLAKIMADIRAQYDELAQRNREELDKYSSQIEESTTWTTSQAEVGAETTLTEL 299
Db 232 PKSQLAKIMADIRAQYDELAQRNREELDKYSSQIEESTTWTTSQAEVGAETTLTEL 291
QY 300 RRTVQSLIEDLSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGQRQ 359
Db 292 RRTVQSLIEDLSMRNLKASLENSLREVEARYALQMEQNGILLHLESELAQTRAEGQRQ 351
QY 360 AQEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDGDLSDSSNSMOTIKTTTRIVDGRV 419
Db 352 AQEYEAALLNIKVKLEAEIATYRRLLEDGEDFNLDGDLSDSSNSMOTIKTTTRIVDGRV 411
QY 420 SETNDTKVLRH 430
Db 412 SETNDTKVLRH 422
```



```

RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; BC042269; AAH42269.1; -.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; UNKNOWN_1.
KW Intermediate filament.
FT NON TER 1
SQ SEQUENCE 436 AA; 48322 MW; F3466A556C5B729D CRC64;

Query Match 59.0%; Score 1250.5; DB 2; Length 436;
Best Local Similarity 59.0%; Pred. No. 4.4e-54;
Matches 260; Conservative 82; Mismatches 68; Indels 31; Gaps 9;

QY 4 TTRSTFTN-----YRSLGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTSPRG 56
DB 11 SRSVYSSSSVVGSPYRSLSS--APRF--TPGSAASVHAGAGSGARISVSRVSTVGS 66
QY 57 GMGSGGLATGATGAGLAG-----MGIGQNEKETMQSLNDRLASVLDVRVSRLETRRLRS 110
DB 67 GFG-----GGFGSASNVNLFQGVQNEKETMQSLNDRLASVLDVRVSRLETRRLRS 117
QY 111 KIREHLEKGPQVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYET 170
DB 118 QIRQHTKKGKGP-AKDWSPYMTIEDLKKQVFNISVENSQVLQIDNARLAADDFRVKYES 176
QY 171 ELAMRQSVENDIHGLRKVIDDNTNITRIQLETEIEALKEELLFMKNHEEVGGLQAOIAS 230
DB 177 EVAIRMSVETDGLGLRLIDDTNISRLNLENFESLKEELIFLKNQHQDDVNELOAOIAS 236
QY 231 SGLTVEVDAPKSQDLAKIMADIRAOYDELARKNRELDKYWSQIEESTVVTQSAEVG 290
DB 237 SAVTVEVDAPKSQDLGKIMADLRAQYDEMAQKNREDVEKLYQSKVEEHTVQVNLDAEALH 296
QY 291 AAEITTLTDLRRTVQSLIEDLDSMRNLKASLNSLREVEARYALQMEQLNGILLHLESELA 350

```

```

DB 297 TAKSSVTELRRTMQSLIELESLRNQKASLEGTLHTDEARYAMELEMDGGTAMALETELV 356
QY 351 QTRAEGQROAQEYEAALLNIKVKLEAEIATYRLLLEDGEDFNLDGALDSNSMQTIQK--T 408
DB 357 QVRNDCQRCQEQYQALLNTKMKLEABIQTYRRLLE-GDSFDLQDAVPVVT--QTVKKVIT 414
QY 409 TTRRIVDGKVVSETNDTKVLR 429
DB 415 TTQRLVDGKVAESNNTTEVIK 435
RESULT 7
Q6P864 PRELIMINARY; PRT; 429 AA.
AC Q6P864;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75922.
GN Name=MGC75922;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; BC061366; AAH61366.1; -.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; UNKNOWN_1.
KW Hypothetical protein; Intermediate filament.
SQ SEQUENCE 429 AA; 47613 MW; 416F307F66D55101 CRC64;

Query Match 58.9%; Score 1249.5; DB 2; Length 429;
Best Local Similarity 58.6%; Pred. No. 4.8e-54;
Matches 260; Conservative 84; Mismatches 69; Indels 31; Gaps 9;

QY 1 MSFTTSTFTSTN-----YRSLGVSQAPSYGARPVSSAASVYAGAGSGSRISVSRSTS 53
DB 1 MSY-SRSVYSSSSVVGSPYRSLSS--APRF--IPSSAASVHAGAGSGARISVSRVSS 55

```

KW Hypothetical protein.
SQ SEQUENCE 429 AA; 47613 MW; 416P307F66D55101 CRC64;
Query Match 58.9%; Score 1249.5; DB 2; Length 429;
Best Local Similarity 58.6%; Pred. No. 4.8e-54;
Matches 260; Conservative 84; Mismatches 69; Indels 31; Gaps 9;
QY 1 MSFTTRSTFTSN-----YRSLGSVQAPSYGARPVSSAASVYAGAGSGSRISVSRSTS 53
DB 1 MSY-SRSVSSSVGGSPYRSLSS--APRF--IPSSAASVHAGAGSGGARISVSRVSS 55
QY 54 FRGMGSGGLATGIAGGLAG-----MGIONEKETWQSLNDRSLASVLDVRSLETENRR 107
DB 56 VSGSGF-----GGYSGVSNVNLIGGQNEKETWQSLNDRSLASVLDVRSLETENRR 106
QY 108 LESKIREHLEKKGPOVRDWSHYFKIIEIDLRQAFANTVDNARIVLQIDNARLAADDPRVK 167
DB 107 LEVQIRQHTKKGPS-KDWSPPYKTIEDLRKQVDFSTLNSQLVLDNARLAADDPRVK 165
QY 168 YETELAMRQSVENDIHGLRKVIDDTNITRLQLETEIEALKEELFMKNHEEVEVKGLOAQ 227
DB 166 YEAEWAIRMSVEGDITGLRKLIDDTNVSRLNENEIESLKEELIFLKNHQDDVTELQAO 225
QY 228 IASSGLTVEVDAPKSDQLAKIMADIRAOYDELARKNREELDKYWSQOIEESTTVTTTQSA 287
DB 226 VARSAVTVEVDAPKSDQLGKIMTELRAQYDGLAQKRDVDEKWKYQSKVDEHTVQVNLDT 285
QY 288 EVGAAETTLTLELRRTVQSLDLSMRNLKASLENSREVEARYALQMEQLNGILLHLES 347
DB 286 ALHSAKSSVTLRRTVQSLDLSMRNLKASLENSREVEARYALQMEQLNGILLHLES 345
QY 348 ELAQTRAEQGRQAOEYEAALLNKKVLEAEIATYRRLLLEDGEDFNLDALDSSNSMQTIQK 407
DB 346 ELVQVRSDCRQOQOYQALLNTKMLEAEIHTYRRLLLEDGEDFNLDALDSSNSMQTIQK 404
QY 408 --TTTTRIVDGKVVSEETNDTKVLR 429
DB 405 VITTTQRIVDGKVVSEETNDTEVLK 428

RESULT 9

K1CR_XENLA
ID K1CR_XENLA STANDARD; PRT; 368 AA.
AC P08802;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Keratin, type I cytoskeletal endo B (fragment).
GN Name=XK ENDO B;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89092007; PubMed=2463213;
RA Laflamme S.E., Jamrich M., Richter K., Sargent T.D., Dawid I.B.;
RT "Xenopus endo B is a keratin preferentially expressed in the embryonic notochord."
RL Genes Dev. 2:853-862(1988).
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and microfibrillar keratin, I (acidic) and II (neutral to basic) (40-55 and 56-70 kDa, respectively).
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

QY 54 FRGMGSGGLATGIAGGLAG-----MGIONEKETWQSLNDRSLASVLDVRSLETENRR 107
DB 56 VSGSGF-----GGYSGVSNVNLIGGQNEKETWQSLNDRSLASVLDVRSLETENRR 106
QY 108 LESKIREHLEKKGPOVRDWSHYFKIIEIDLRQAFANTVDNARIVLQIDNARLAADDPRVK 167
DB 107 LEVQIRQHTKKGPS-KDWSPPYKTIEDLRKQVDFSTLNSQLVLDNARLAADDPRVK 165
QY 168 YETELAMRQSVENDIHGLRKVIDDTNITRLQLETEIEALKEELFMKNHEEVEVKGLOAQ 227
DB 166 YEAEWAIRMSVEGDITGLRKLIDDTNVSRLNENEIESLKEELIFLKNHQDDVTELQAO 225
QY 228 IASSGLTVEVDAPKSDQLAKIMADIRAOYDELARKNREELDKYWSQOIEESTTVTTTQSA 287
DB 226 VARSAVTVEVDAPKSDQLGKIMTELRAQYDGLAQKRDVDEKWKYQSKVDEHTVQVNLDT 285
QY 288 EVGAAETTLTLELRRTVQSLDLSMRNLKASLENSREVEARYALQMEQLNGILLHLES 347
DB 286 ALHSAKSSVTLRRTVQSLDLSMRNLKASLENSREVEARYALQMEQLNGILLHLES 345
QY 348 ELAQTRAEQGRQAOEYEAALLNKKVLEAEIATYRRLLLEDGEDFNLDALDSSNSMQTIQK 407
DB 346 ELVQVRSDCRQOQOYQALLNTKMLEAEIHTYRRLLLEDGEDFNLDALDSSNSMQTIQK 404
QY 408 --TTTTRIVDGKVVSEETNDTKVLR 429
DB 405 VITTTQRIVDGKVVSEETNDTEVLK 428
RESULT 8
AAH61366
ID AAH61366 PRELIMINARY; PRT; 429 AA.
AC AAH61366;
DT 25-MAR-2004 (TREMBlrel. 27, Created)
DT 25-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 25-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein MGC75922.
GN MGC75922.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061366; AAH61366.1; --

CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Y00230; CAAG8372.1; -.
DR PIR; A28825; A28825.
DR HSSP; P08670; 1GK7.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; 1.
KW Coiled coil; Intermediate filament; Keratin.
FT NON TER 1 15 Head.
FT DOMAIN <1 15 Rod.
FT DOMAIN 16 325 Rod.
FT DOMAIN 326 368 Tail.
FT DOMAIN 16 54 Coil 1A.
FT DOMAIN 55 66 Linker 1.
FT DOMAIN 67 162 Coil 1B.
FT DOMAIN 163 183 Linker 2.
FT DOMAIN 184 325 Coil 2.
SQ SEQUENCE 368 AA; 41683 MW; 29FC19B67E9E8F95 CRC64;

Query Match 53.7%; Score 1137.5; DB 1; Length 368;
Best Local Similarity 61.2%; Pred. No. 1.4e-48;
Matches 226; Conservative 74; Mismatches 58; Indels 11; Gaps 5;

QY 69 GGLAG-----MGIONEKETMQSLNDRLASVLDVRSLENTNRRLSKIREHLEKGPQ 122
Db 2 GGFSGASNNVLPFGVQNEKETMQDLNDRLASVLDVRSLENTNRRLSKIREHLEKGP- 60

QY 123 VDWSHYFKLIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVENDI 182
Db 61 AKDWSPYMTIEDLKKQVFNISVENSQVLDVNDARLAADDFRVKYETELAMRQSVETDI 120

QY 183 HGLRKVIDDNTNITRLQLETEIEALKKEELLMKKNHEEVEKGLQAIASSGLTVEVDAPKS 242
Db 121 GGLRKLIDNTNISRLNLENFESLKEELIFLKNHQDDVNEIQAQIATSATVVEVDAPKS 180

QY 243 QDLAKIMADIRAOYDELARKNREELDKYWSQIIESTTVVTTQSAEVGAAETTLTELRR 302
Db 181 QDLGKIMADIRAOYDEMAQKNREDVEKLYQSKVEDDTQVNLDAEALHTAKSSVTELRR 240

QY 303 VQSLIEDLDSMRNLKASLENSREVEARYALQMEQNGILLHLESELAQTRAEGQRAQE 362
Db 241 MQSLEIELESLRNQKASLEGTLHDTFARYAMELEMLGGTAMALETELQVRNDCQRCQOE 300

QY 363 YEALLNKKVLEAEIATYRRLLLEDGEDFNGLDSSNSMOTIQK--TTTTRIVDGKVS 420
Db 301 YQALLNKKVLEAEIATYRRLLLEDGEDFNGLDSSNSMOTIQK--TTTTRIVDGKVS 358

QY 421 ETNDTKVLR 429
Db 359 ESNNTTEVIK 367

RESULT 10
Q7SYF8 PRELIMINARY; PRT; 435 AA.
AC Q7SYF8;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Keratin type is.
GN Name=krt18;
OS Acipenser baerii (Siberian sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=27689;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mixture of gill;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; AJ493261; CAD38124.1; -.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Keratin.
SQ SEQUENCE 435 AA; 48486 MW; 2DADCEFI84C9E563 CRC64;

Query Match 53.6%; Score 1135.5; DB 2; Length 435;
Best Local Similarity 55.6%; Pred. No. 2.1e-48;
Matches 240; Conservative 83; Mismatches 102; Indels 7; Gaps 4;

QY 1 MSFTTRSTFTNYSRLGVSQAPS---YGARPVSSAASVYAGAGGSGSRISVSRTS-FR 55
Db 1 MSYRPGSYSVSSMRPVGSVRSQVMTVQSRMPLASAAVYGGAGGSRISVGGSSGFG 60

QY 56 GGMGSGGLATGIAGGLAGMGIGIONEKETMQSLNDRLASVLDVRSLENTNRRLSKIREH 115
Db 61 SGLSGAGGSSYSSMSVSGSLVGNKEKETMIGLNDRLAAYLETVRNLEQANSKLEFQIREA 120

QY 116 LEKKGQVDRDWSHYFKLIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMR 175
Db 121 LEKKGQVDRDWSHYFKLIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMR 180

QY 176 QSVENDIHGLRKVIDDNTNITRLQLETEIEALKKEELLMKKNHEEVEKGLQAIASSGLT 235
Db 181 QSVESDIIGLRKVIDDNTNITRLQLETEIEALKKEELLMKKNHEEVEKGLQAIASSGLT 240

QY 236 EVDAPKSDLAKIMADIRAOYDELARKNREELDKYWSQIIESTTVVTTQSAEVGAAET 295
Db 241 DVDAPKSDLAKIMADIRAOYDELARKNREELDKYWSQIIESTTVVTTQSAEVGAAET 300

QY 296 LTELRTVQSLIEDLDSMRNLKASLENSREVEARYALQMEQNGILLHLESELAQTRA 355
Db 301 VTELRRQMSLEIELESQSRMSKASLEDSLRDTEMNRNMEMERYNNMILQLEAEELQRLGN 360

QY 356 GORQAEYEALLNKKVLEAEIATYRRLLLEDGEDFNGLDSSNSMOTIQK--TTTTRIV 414
Db 361 IQMQAEYEALLNKKVLEAEIATYRRLLLEDGEDFNGLDSSNSMOTIQK--TTTTRIV 419

QY 415 DGKVVSETNDTK 426
Db 420 DGKVVSESTNTK 431

RESULT 11
Q7ZTS4 PRELIMINARY; PRT; 431 AA.
AC Q7ZTS4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Keratin 18.
GN Name=krt18;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

QY 300 RRTVQSLEIDLSMRNLKASLENSLREVEARYALOWEQLNGILLHLESELAQTRAEGRQ 359
 DB 304 RQIQTLEIELESOKNLKSGLEGLTRDTMYNMEIENLTIIQLAEALTQLRGNIQHQ 363
 QY 360 AQEYEALLNIKVKLEAEIATYRRLLLEDGEDFNGLDSSNSMOTIQKTTTTRIVDGKVV 419
 DB 364 TOEYEALLNIKVKLEAEIATYRRLL-DGGDFKLQDALEEQKVKVM--TVTQTLVDGKVV 420
 QY 420 SETNDTK 426
 DB 421 SSSSTETK 427

RESULT 13
 AAH65848
 ID AAH65848 PRELIMINARY; PRT; 431 AA.
 AC AAH65848;
 DT 02-MAR-2004 (TREMBLrel. 27, Created)
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
 DE Keratin 18.
 GN KRT18.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Strausberg R.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC065848; AAH65848.1; -.
 KW Keratin.
 SQ SEQUENCE 431 AA; 48631 MW; 851BDACDFC8996B7 CRC64;

Query Match 51.5%; Score 1092.5; DB 2; Length 431;
 Best Local Similarity 54.1%; Pred. No. 2.7e-46;
 Matches 231; Conservative 74; Mismatches 115; Indels 7; Gaps 4;
 QY 2 SFTTRSTFTNYSGLSVQAPSYGAPVSSAASVYAGAGGSGRISVSRSTSFRCGMSG 61
 DB 6 SYSVRS--STSQVPVSQVSIKRTTNTVPTYRAASVYAGAGGSGRISVSRSTSFRCGMSG 63
 QY 62 GLATGIAGGLAGMG--IQNEKETMQLNDRLASVLDVRSLETENRRLESKIREHLEKK 119
 DB 64 SMSSSIQVSASGSTGEIMGNEKAMQNLNDRLASVLDVRSLETENRRLESKIREHLEKK 123

QY 120 GPOVRDWSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKXETELAMEQSV 179
 DB 124 GPDVHDYSRFQPIVDLKKIFDATTNNARLVQIDNARLAADDFRVKXETELAMEQSV 183
 QY 180 NDIHGLRKVIDDTNITRLQLETEALKEELFMKKNHEEVKGLQAOIASSGLTVEYDA 239
 DB 184 ADITGLRKVIDDTNLRNMNLESETEALKEELFLKKNHNEVMELRNQISQSGVQVDVA 243
 QY 240 PKSQLAKIMADIRAOYDELARKNREELDQYMSQIESTTWTTSAAEVGAAETTLTEL 299
 DB 244 PKGQDLSQIMEIRAKYEMALKNQEELKAWHESQITEVQVQVQNTQNTALQGARSEVNE 303
 QY 300 RRTVQSLEIDLSMRNLKASLENSLREVEARYALOWEQLNGILLHLESELAQTRAEGRQ 359
 DB 304 RQIQTLEIELESOKNLKSGLEGLTRDTMYNMEIENLTIIQLAEALTQLRGNIQHQ 363
 QY 360 AQEYEALLNIKVKLEAEIATYRRLLLEDGEDFNGLDSSNSMOTIQKTTTTRIVDGKVV 419
 DB 364 TOEYEALLNIKVKLEAEIATYRRLL-DGGDFKLQDALEEQKVKVM--TVTQTLVDGKVV 420
 QY 420 SETNDTK 426
 DB 421 SSSSTETK 427

RESULT 14
 OS7607
 ID OS7607 PRELIMINARY; PRT; 438 AA.
 AC OS7607;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE K18, simple type I keratin.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=99054191; PubMed=9840456;
 RA Schaffeld M., Loebbecke A.B., Lieb B., Markl J.;
 RT "Tracing keratin evolution: catalog, expression patterns and primary
 RT structure of shark (Scyliorhinus stellaris) keratins."
 RL Eur. J. Cell Biol. 77:69-80(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Conrad M.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the intermediate filament family.
 DR EMBL; Y14289; CAA74664.1; -.
 DR HSSP; P08670; 1CK7.
 DR GO; GO:0005882; C:intermediate filament; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR002957; Keratin_I.
 DR Pfam; PF00038; Filament; 1.
 DR PRINTS; PR01248; TYPE1KERATIN.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament.
 SQ SEQUENCE 438 AA; 48779 MW; 06EF8C4EEB841B CRC64;

Query Match 51.1%; Score 1082.5; DB 2; Length 438;
 Best Local Similarity 54.0%; Pred. No. 8.7e-46;
 Matches 232; Conservative 68; Mismatches 107; Indels 23; Gaps 4;
 QY 18 SVQAPSYG-----ARPVSSAASVYAGAGGSGRISVSRSTSFRCGMSGGLATGIA 68
 DB 7 SVRGPSGSGYGSITHNSTAPTTRAASTYGGAGGQGTTRISVSVSGVSRGMSGVGMGGS 66
 QY 69 GGLAGMGGIQ-----NEKETMQLNDRLASVLDVRSLETENRRLESKIREHL 116

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 09:19:20 ; Search time 192 Seconds
(without alignments)
1288.599 Million cell updates/sec

Title: US-10-026-001-1
Perfect score: 2120
Sequence: 1 MSFTTSTSTNYRSLGSVQ.....RRIVDGKVVSETNDTKVLRH 430

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 30763

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	4.2	21	Q9NPD2	Q9npd2 homo sapien
2	82	3.9	28	Q9QWC0	Q9qwc0 rattus sp.
3	58.5	2.8	23	Q28131	Q28131 bos taurus
4	58	2.7	16	Q9UCY4	Q9ucy4 homo sapien
5	49	2.3	24	Q9R5X7	Q9r5x7 mycoplasma.
6	48.5	2.3	27	Q9QV51	Q9qv51 mus sp. 60
7	46	2.2	17	Q9PS62	Q9ps62 notophthalm
8	44	2.1	18	Q9PS62	Q9ps62 physarum po
9	43	2.0	30	Q81SU8	Q81su8 bacillus an
10	43	2.0	30	AAT30644	Aat30644 bacillus
11	42	2.0	23	Q86MM6	Q86mm6 pecten maxi
12	42	2.0	25	Q42086	Q42086 arabidopsis
13	41.5	2.0	27	Q9QHC5	Q9qhc5 hepatitis c
14	41	1.9	29	Q6V4L2	Q6v4l2 streptococ
15	41	1.9	29	AAQ64524	AAq64524 streptoco
16	40.5	1.9	27	Q9QHC2	Q9qhc2 hepatitis c
17	40.5	1.9	27	Q9QHC8	Q9qhc8 hepatitis c
18	40.5	1.9	27	Q9QHC9	Q9qhc9 hepatitis c
19	40.5	1.9	27	Q9QHD4	Q9qhd4 hepatitis c
20	40	1.9	22	Q9MX47	Q9mx47 oryzias lat
21	40	1.9	22	Q94R02	Q94r02 clarias bat
22	39.5	1.9	26	Q42226	Q42226 arabidopsis
23	39	1.8	12	Q6WR52	Q6wr52 asio otus
24	39	1.8	12	Q6WR73	Q6wr73 coccyzus er
25	39	1.8	12	AAQ23430	AAq23430 coccyzus
26	39	1.8	12	AAQ23451	AAq23451 asio otus
27	39	1.8	25	Q9TRW8	Q9trw8 trichosurus
28	39	1.8	28	Q37931	Q37931 bacterioph
29	39	1.8	29	Q718V7	Q718v7 newcastle d
30	39	1.8	29	AAQ11541	AAq11541 newcastle
31	39	1.8	30	Q6RBX4	Q6rbx4 homo sapien

32	39	1.8	30	2	Q6RBR0	Q6rbr0 pygathrix b
33	39	1.8	30	2	Q6RBR1	Q6rbr1 hyllobates l
34	39	1.8	30	2	Q6RBR2	Q6rbr2 hyllobates h
35	39	1.8	30	2	Q6RBR4	Q6rbr4 macaca mula
36	39	1.8	30	2	Q6RBR5	Q6rbr5 erythrocebu
37	39	1.8	30	2	Q6RBR6	Q6rbr6 trachypithe
38	39	1.8	30	2	Q6RBR7	Q6rbr7 pygathrix n
39	39	1.8	30	2	Q6RBR8	Q6rbr8 trachypithe
40	39	1.8	30	2	Q6RBS1	Q6rbs1 pongo pygma
41	39	1.8	30	2	Q6RBS5	Q6rbs5 gorilla gor
42	39	1.8	30	2	AAS86449	Aas86449 homo sapi
43	39	1.8	30	2	AAS86450	Aas86450 homo sapi
44	39	1.8	30	2	AAS86451	Aas86451 homo sapi
45	39	1.8	30	2	AAS86452	Aas86452 homo sapi

ALIGNMENTS

RESULT 1
Q9NPD2
ID Q9NPD2 PRELIMINARY; PRT; 21 AA.
AC Q9NPD2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Keratin 19 (Fragment).
GN Name=keratin 19;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato T., Weerasinghe A., Kuwano Y., Kaneko T., Ikeda T., Nagai T.,
RA Makino H., Sano M., Honma K., Nemoto K., Abo T., Shima Y.;
RT "Diversity of keratin 19 gene expressed in lymph nodes of breast cancer
patients - Strategy to clear the discrepancy between histological
findings and RT-PCR results in the detection of micrometastasis."
RL Seibutsu Butsuri Kagaku 44:201-204(2000).
DR EMBL; AB041270; BAA94610.1; -
DR EMBL; AB041269; BAA94609.1; -
DR GO; GO:0005882; C:intermediate filament; IEA.
KW Keratin.
FT NON TER 1
SQ SEQUENCE 21 AA; 2330 MW; 7AFBDA77D61F22B7 CRC64;
Query Match 4.2%; Score 88; DB 2; Length 21;
Best Local Similarity 81.0%; Pred. No. 3.5e+02;
Matches 17; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 144 TVDNARIVLQIDNARLAADD 164
Db 1 TIENARIVLQINNAQLAADD 21
RESULT 2
Q9QWC0
ID Q9QWC0 PRELIMINARY; PRT; 28 AA.
AC Q9QWC0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Cytokeratin D, P45 (Fragments).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93077141; PubMed=1280245;
RA Bastos R., Engel P., Pujades C., Falchetto R., Alique R., Bachs O.;
RT "Increase of cytokeratin D during liver regeneration: association with


```

RT the nuclear matrix."
RL Hepatology 16:1434-1446(1992).
FT NON_TER 1
FT NON_CONS 7
FT NON_TER 28
SQ SEQUENCE 28 AA; 3078 MW; 18B27A72DC43703E CRC64;

Query Match 3.9%; Score 82; DB 2; Length 28;
Best Local Similarity 78.3%; Pred. No. 9.5e+02;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 276 EESTVVTTCQAEVGAETTLTE 298
Db 6 EESTVVTTKSABIRDAETTLLE 28

RESULT 3
Q28131 ID Q28131 PRELIMINARY; PRT; 23 AA.
AC Q28131;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Epidermal keratin type II IA (Fragment).
GN Name=keratin type II IA;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=85126936; PubMed=6084595;
RA Lehnert M.E., Jorcano J.L., Zentgraf H., Blessing M., Franz J.K.,
RA Franke W.W.;
RT "Characterization of bovine keratin genes: similarities of exon
RT patterns in genes coding for different keratins.";
RL EMBO J. 3:3279-3287(1984).
DR EMBL; M28275; AAA30600.1; -.
DR PIR; I45916; I45916.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
KW Keratin.
FT NON_TER 1
FT NON_TER 23
SQ SEQUENCE 23 AA; 2566 MW; BC381E2C61B3EE59 CRC64;

Query Match 2.8%; Score 58.5; DB 2; Length 23;
Best Local Similarity 56.5%; Pred. No. 1.1e+04;
Matches 13; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 367 LNIKVKLEAEIATYRRLLEDGED 389
Db 1 MNVKLALDIEIATYPTLLE-GE 22

RESULT 4
Q9UCY4 ID Q9UCY4 PRELIMINARY; PRT; 16 AA.
AC Q9UCY4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Keratin 14 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=96179748; PubMed=8601736;
RA Chan Y.M., Cheng J., Gedde-Dahl T.Jr., Niemi K.M., Fuchs E.;
RT "Genetic analysis of a severe case of Dowling-Meara epidermolysis
RT bullosa simplex.";
RL J. Invest. Dermatol. 106:327-334(1996).
SQ SEQUENCE 16 AA; 1833 MW; BDI20FOA07B7AE72 CRC64;

Query Match 2.7%; Score 58; DB 2; Length 16;
Best Local Similarity 80.0%; Pred. No. 7.7e+03;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 83 TMOQLNDRLASLYLDR 97
Db 2 TMOQLNDRLASDLCK 16

RESULT 5
Q9RSX7 ID Q9RSX7 PRELIMINARY; PRT; 24 AA.
AC Q9RSX7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Epidermal keratin type I intermediate filament protein homolog
DE (Fragment).
OS Mycoplasma.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae.
OX NCBI_TaxID=2093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94174734; PubMed=8128634;
RA Turk S.M., Hutt-Fletcher L.M.;
RT "The Epstein-Barr virus-associated protein p105 is not encoded by the
RT Epstein-Barr virus genome.";
RL Virology 200:313-318(1994).
SQ SEQUENCE 24 AA; 2940 MW; 4E5254E5661847F7 CRC64;

Query Match 2.3%; Score 49; DB 2; Length 24;
Best Local Similarity 33.3%; Pred. No. 3.4e+04;
Matches 8; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 252 IRAQYDELARKNREELDKYWSQI 275
Db 1 MRQEYEQIARKRKDIEGQYEQI 24

RESULT 6
FBRL_PHYPO ID FBRL_PHYPO STANDARD; PRT; 27 AA.
AC P22508;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fibrillarin (34 kDa nucleolar protein B-36) (Fragment).
OS Physarum polycephalum (Slime mold).
OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;
OC Physarum.
OX NCBI_TaxID=5791;
RN [1]
RP SEQUENCE.
RX MEDLINE=89025797; PubMed=3140806;
RA Christensen M.E., Fuxa K.P.;
RT "The nucleolar protein, B-36, contains a glycine and dimethylarginine-
RT rich sequence conserved in several other nuclear RNA-binding
RT proteins.";
RL Biochem. Biophys. Res. Commun. 155:1278-1283(1988).
CC -!- FUNCTION: Fibrillarin is a component of a nucleolar small nuclear
CC ribonucleoprotein particle thought to participate in the first
CC step in processing preribosomal RNA. It is associated with the U3,
CC U8 and U13 small nuclear RNAs.
CC -!- SUBCELLULAR LOCATION: Nuclear; fibrillar region of the nucleolus.
CC -!- SIMILARITY: Belongs to the fibrillarin family.
DR PIR; A31508; A31508.

```

DR InterPro; IPR000692; Fibrillarin.
DR PROSITE; PS00566; FIBRILLARIN; PARTIAL.
KW Direct protein sequencing; Methylation; Nuclear protein;
KW Ribonucleoprotein; RNA-binding; rRNA processing.
FT MOD_RES 5 Asymmetric dimethylarginine.
FT MOD_RES 11 11 Asymmetric dimethylarginine.
FT MOD_RES 16 16 Asymmetric dimethylarginine.
FT MOD_RES 19 19 Asymmetric dimethylarginine.
FT NON_TER 27
SQ SEQUENCE 27 AA; 2462 MW; F76AD7F8FAF442DA CRC64;

Query Match 2.3%; Score 48.5; DB 1; Length 27;
Best Local Similarity 38.7%; Pred. No. 4.1e+04;
Matches 12; Conservative 1; Mismatches 13; Indels 5; Gaps 1;

QY 36 YAGAGGSGSRISVSSTSPRGGMGSGGLATG 66
DB 2 FEGRGGFG-----GRGGGDRGGXGGFGGG 27

RESULT 7
Q9QV51 ID Q9QV51 PRELIMINARY; PRT; 17 AA.
AC Q9QV51;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 60 kDa vimentin homolog (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=94039737; PubMed=7693509;
RA Brandes R., Arad R., Gaathon A., Bar-Tana J.;
RT "Induction of adipose conversion in 3T3-L1 cells is associated with an
RT early phosphorylation of a protein partly homologous with mouse
RT vimentin.";
RL FEBS Lett. 333:179-182(1993).
FT NON_TER 1 1
FT NON_TER 17 17
FT SEQUENCE 17 AA; 2158 MW; 63A1DC9232086211 CRC64;

Query Match 2.2%; Score 46; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 3.2e+04;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 84 MQSLNDRLASYLDRVR 99
DB 1 LQELNDAFMRYDKVR 16

RESULT 8
Q9PS62 ID Q9PS62 PRELIMINARY; PRT; 18 AA.
AC Q9PS62;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 55 kDa XENOPUS keratin B2 homolog (Fragment).
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC Notophthalmus.
OX NCBI_TaxID=8316;
RN [1]
RP SEQUENCE.
RX MEDLINE=92162009; PubMed=1371385;
RA Tsonis P.A., Mescher A.L., Del Rio-Tsonis K.;
RT "Protein synthesis in the newt regenerating limb. Comparative two-
RT dimensional PAGE, computer analysis and protein sequencing.";
RL Biochem. J. 281:665-668(1992).

DR PIR; A56791; A56791.
FT NON_TER 1 1
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1970 MW; 2477C624B2012E8C CRC64;

Query Match 2.1%; Score 44; DB 2; Length 18;
Best Local Similarity 61.5%; Pred. No. 4.3e+04;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 244 DLAKIMADIRAOY 256
DB 4 DLGXILSDMRAOY 16

RESULT 9
Q81SUS ID Q81SUS PRELIMINARY; PRT; 30 AA.
AC Q81SUS; Q6KUY1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BA1547; ORFNames=GBAA1547;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.C., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaiter J.E., White O.,
RA Salzberg S.L., Thomason B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RX Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017029; AAP25483.1; -.
DR EMBL; AE017334; AAT30644.1; -.
DR TIGR; BA1547; -.
KW Hypothetical protein.
SQ SEQUENCE 30 AA; 3613 MW; 8B465215541A80CD CRC64;

Query Match 2.0%; Score 43; DB 2; Length 30;
Best Local Similarity 42.1%; Pred. No. 8.5e+04;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 206 LKEELLFMKKHHEEVKGL 224
DB 1 MKESILYNNRVELEERIKGV 19

RESULT 10
AAT30644 ID AAT30644 PRELIMINARY; PRT; 30 AA.
AC AAT30644;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)

DT 01-JUN-2004 (Tremblrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN GBAAL547.
 OS Bacillus anthracis str. Ames 0581.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
 OC Bacillus cereus group; Bacillus anthracis.
 OX NCBI_TaxID=261594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames 0581;
 RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
 RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
 RT "Bacillus anthracis comparative genomics."
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017334; AAT30644.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 30 AA; 3613 MW; 8E465215541A80CD CRC64;

Query Match 2.0%; Score 43; DB 2; Length 30;
 Best Local Similarity 42.1%; Pred. No. 8.5e+04;
 Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 206 LKEELLFMKNHEEVKGL 224
 Db 1 MKESLLYNRVELEIRKGV 19

RESULT 11

Q86MM6 PRELIMINARY; PRT; 23 AA.
 ID Q86MM6;
 AC Q86MM6;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Cdx homeodomain protein (Fragment).
 GN Name=cdx;
 OS Pecten maximus (King scallop) (Pilgrim's clam).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
 OC Pectinoidea; Pectinidae; Pecten.
 OX NCBI_TaxID=6579;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22967154; PubMed=14604796;
 RA Barucca M., Olmo E., Canapa A.;
 RT "Hox and paraHox genes in bivalve molluscs."
 RL Gene 317:97-102(2003).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 DR EMBL; AJ344456; CAD58905.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD00010; Homeobox; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2783 MW; AB91DF5413EEBS17 CRC64;

Query Match 2.0%; Score 42; DB 2; Length 23;
 Best Local Similarity 52.6%; Pred. No. 7e+04;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 341 ILLHLESELACTRAEGORQ 359
 Db 5 ITIRKSELATLALSERQ 23

RESULT 12

Q42086 PRELIMINARY; PRT; 25 AA.
 ID Q42086
 AC Q42086;
 DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Glycine-rich RNA binding protein (Fragment).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cell suspension culture of A.thaliana ecotype;
 RA CNRS;
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cell suspension culture of A.thaliana ecotype;
 RA Philipps G., Gigot C.;
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z26410; CAA81244.1; -.
 FT NON_TER 1
 SQ SEQUENCE 25 AA; 2421 MW; E1064170E1E33751 CRC64;

Query Match 2.0%; Score 42; DB 2; Length 25;
 Best Local Similarity 55.6%; Pred. No. 7.7e+04;
 Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 59 GSGGLATGIAGGLAGMG 76
 Db 7 GSGGGGGYGGGGGGGG 24

RESULT 13

Q9QHC5 PRELIMINARY; PRT; 27 AA.
 ID Q9QHC5
 AC Q9QHC5;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Izopet J.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF167003; AAD53636.1; -.
 KW Polyprotein.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2565 MW; C46288D07BF3C2D0 CRC64;

Query Match 2.0%; Score 41.5; DB 2; Length 27;
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;
 Matches 11; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 61 GGLATGIAGGLAGM--GGIQN 79
 Db 6 GGIAGRAAAGLAGLFRPGSQN 27

RESULT 14

Q6V4L2 PRELIMINARY; PRT; 29 AA.
 ID Q6V4L2
 AC Q6V4L2;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE M protein (Fragment).
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.

OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS223;
RX PubMed=1468117;
RA McKay F.C., McArthur J.D., Sanderson-Smith M.L., Gardam S.,
RA Currie B.J., Sriprakash K.S., Fagan P.K., Towers R.J., Batzloff M.R.,
RA Chhatwal G.S., Ranson M., Walker M.J.;
RT "Plasminogen Binding by Group A streptococcal isolates from a region
RT of hyperendemicity for streptococcal skin infection and a high
RT incidence of invasive infection."
RL Infect. Immun. 72:364-370(2004).
DR EMBL; AY351854; AAQ64524.1; -.
FT NON_TER 1 29
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3697 MW; C22EE2EC8C34E229 CRC64;

Query Match 1.9%; Score 41; DB 2; Length 29;
Best Local Similarity 39.1%; Pred. No. 1e+05;
Matches 9; Conservative 5; Mismatches 5; Indels 4; Gaps 1;

QY 200 ETEIEALKKEELLFMKKNHEEVK 222
| | | | | : : : : :
Db 1 EVELERLKNE----RHDHDEAE 19

RESULT 15
AAQ64524
ID AAQ64524 PRELIMINARY; PRT; 29 AA.
AC AAQ64524;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE M protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS223;
RX PubMed=1468117;
RA McKay F.C., McArthur J.D., Sanderson-Smith M.L., Gardam S.,
RA Currie B.J., Sriprakash K.S., Fagan P.K., Towers R.J., Batzloff M.R.,
RA Chhatwal G.S., Ranson M., Walker M.J.;
RT "Plasminogen Binding by Group A Streptococcal Isolates from a Region
RT of Hyperendemicity for Streptococcal Skin Infection and a High
RT Incidence of Invasive Infection."
RL Infect. Immun. 72:364-370(2004).
DR EMBL; AY351854; AAQ64524.1; -.
FT NON_TER 1 29
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3697 MW; C22EE2EC8C34E229 CRC64;

Query Match 1.9%; Score 41; DB 2; Length 29;
Best Local Similarity 39.1%; Pred. No. 1e+05;
Matches 9; Conservative 5; Mismatches 5; Indels 4; Gaps 1;

QY 200 ETEIEALKKEELLFMKKNHEEVK 222
| | | | | : : : : :
Db 1 EVELERLKNE----RHDHDEAE 19

This Page Blank (uspto)

A31508

34K nucleolar protein B-36 - slime mold (Physarum polycephalum) (fragment)
C:Species: Physarum polycephalum
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: A31508

R;Christensen, M.E.; Fuxa, K.P.
Biochem. Biophys. Res. Commun. 155, 1278-1283, 1988
A:Title: The nucleolar protein, B-36, contains a glycine and dimethylarginine-rich sequence
A:Reference number: A31508; MUID:89025797; PMID:3140806
A:Accession: A31508
A:Molecule type: protein
A:Residues: 1-27 <CHR>
A:Cross-references: UNIPROT:P22508
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: methylated amino acid; nucleolus
F:5,11,16,19/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status experimental

Query Match	2.3%;	Score 48.5;	DB 2;	Length 27;
Best Local Similarity	38.7%;	Pred. No. 5e+03;		
Matches	12;	Conservative	1;	Mismatches 13; Indels 5; Gaps 1;

```

QY      36 YAGAGGGSGRISVSRTSPFRGGMGSGGLATG 66
          :| || |         |||| |   |
Db       2 FEGRGGFG-----GRGGDRGRGXGGFGGG 27

```

RESULT 4
A56791
keratin B2 homolog, 55K - eastern newt (fragment)
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: A56791
R;Tsonis, P.A.; Mescher, A.L.; Del Rio-Tsonis, K.
Biochem. J. 281, 665-668, 1992
A:Title: Protein synthesis in the newt regenerating limb. Comparative two-dimensional PA
A:Reference number: A56791; MUID:92162009; PMID:1371385
A:Accession: A56791
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-18 <TSO>
A:Cross-references: UNIPROT:Q9PS62
A:Experimental source: regenerating limb
A>Note: sequence extracted from NCBI backbone (NCBIP:82650)

Query Match	2.1%;	Score 44;	DB 2;	Length 18;
Best Local Similarity	61.5%;	Pred. No. 5.1e+03;		
Matches	8;	Conservative	3;	Mismatches 2; Indels 0; Gaps 0;

```

QY      244 DLAKIMADIRAQY 256
          || |:::|||
Db       4 DLGXILSDMRAQY 16

```

RESULT 5
A24802
cuticle protein 7 - migratory locust (fragment)
C:Species: Locusta migratoria (migratory locust)
C>Date: 22-Jul-1987 #sequence_revision 22-Jul-1987 #text_change 09-Jul-2004
C:Accession: A24802
R;Hojrup, P.; Andersen, S.O.; Roepstorff, P.
Eur. J. Biochem. 154, 153-159, 1986
A:Reference number: A91157; MUID:86108304; PMID:3943519
A:Accession: A24802
A:Molecule type: protein
A:Residues: 1-24 <HOJ>
A:Cross-references: UNIPROT:P11733

Query Match	1.9%;	Score 41;	DB 2;	Length 24;
Best Local Similarity	45.0%;	Pred. No. 1e+04;		
Matches	9;	Conservative	3;	Mismatches 8; Indels 0; Gaps 0;

```

QY      15 SLGSVQAPSYGARPVSSAAS 34
          |||||:|:|:|:|
Db       4 SAGYAAPPAYAAAPAYAAA 23

```


RESULT 14

S06854
chorion class B protein (variant j7) - silkworm (fragment)
C:Species: Bombyx mori (silkworm)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 12-Apr-1995
C:Accession: S06854
R:Tsitolou, S.G.; Rodakis, G.C.; Alexopoulos, M.; Kafatos, F.C.; Ito, K.; Iatrou, K.
EMBO J. 2, 1845-1852, 1983
A:Title: Structural features of B family chorion sequences in the silkworm Bombyx mori,
A:Reference number: S01420; MUID:84057707; PMID:6571700
A:Accession: S06854
A:Molecule type: protein
A:Residues: 1-29 <TSI>
C:Superfamily: chorion class A protein pc292

Query Match 1.7%; Score 36.5; DB 2; Length 29;
Best Local Similarity 38.5%; Pred. No. 2e+04;
Matches 10; Conservative 0; Mismatches 5; Indels 11; Gaps 1;

QY 37 AGAGGSGSRISVSRSTSRGGMGG 62
|||
Db 6 AGLGGCG-----GGRGYGG 20

RESULT 15

S16336
beta-conglycinin beta chain - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 21-Nov-1993 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: S16336
R:Coates, J.B.; Medeiros, J.S.; Thanh, V.H.; Nielsen, N.C.
Arch. Biochem. Biophys. 243, 184-194, 1985
A:Title: Characterization of the subunits of beta-conglycinin.
A:Reference number: S16334; MUID:86049421; PMID:3840670
A:Accession: S16336
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-26 <COA>
A:Cross-references: UNIPROT:Q22121; UNIPROT:Q93VL9
C:Superfamily: glycinin

Query Match 1.7%; Score 36; DB 2; Length 26;
Best Local Similarity 55.6%; Pred. No. 1.9e+04;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 388 EDFNLGDALDSSNSMOTI 405
|||
Db 5 EDENPFYLRSSNSFQTL 22

Search completed: December 14, 2004, 09:33:40
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 09:27:11 ; Search time 143 Seconds
(without alignments)
1074.033 Million cell updates/sec

Title: US-10-026-001-1
Perfect score: 2120
Sequence: 1 MSFTTRSTFTSTNYSRLGSVQ.....RRIVDGKVVSEITNDTKVLRH 430

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 330368

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	103	4.9	20	9 US-09-915-306-2	Sequence 2, Appli
2	103	4.9	20	9 US-09-915-374-2	Sequence 2, Appli
3	92	4.3	18	9 US-09-915-306-3	Sequence 3, Appli
4	92	4.3	18	9 US-09-915-374-3	Sequence 3, Appli
5	78	3.7	15	9 US-09-915-306-11	Sequence 11, Appl
6	78	3.7	15	9 US-09-915-374-11	Sequence 11, Appl
7	70	3.3	13	9 US-09-915-306-13	Sequence 13, Appl
8	70	3.3	13	9 US-09-915-374-13	Sequence 13, Appl
9	68	3.2	13	9 US-09-915-306-10	Sequence 10, Appl
10	68	3.2	13	9 US-09-915-374-10	Sequence 10, Appl
11	65	3.1	12	9 US-09-915-306-12	Sequence 12, Appl
12	65	3.1	12	9 US-09-915-374-12	Sequence 12, Appl
13	64	3.0	13	10 US-09-988-493-57	Sequence 57, Appl

14	64	3.0	13	17	US-10-700-330-215	Sequence 215, Appl
15	61	2.9	13	10	US-09-988-493-59	Sequence 59, Appl
16	61	2.9	13	17	US-10-700-330-240	Sequence 240, Appl
17	61	2.9	16	17	US-10-774-928-14	Sequence 14, Appl
18	61	2.9	27	14	US-10-352-786-2655	Sequence 2655, Ap
19	61	2.9	28	14	US-10-352-786-2831	Sequence 2831, Ap
20	61	2.9	29	14	US-10-352-786-3015	Sequence 3015, Ap
21	61	2.9	30	14	US-10-352-786-3207	Sequence 3207, Ap
22	60.5	2.9	30	10	US-09-833-203-19	Sequence 19, Appl
23	60	2.8	27	14	US-10-352-786-2653	Sequence 2653, Ap
24	60	2.8	28	14	US-10-352-786-2829	Sequence 2829, Ap
25	60	2.8	29	14	US-10-352-786-3013	Sequence 3013, Ap
26	60	2.8	30	14	US-10-352-786-3205	Sequence 3205, Ap
27	59	2.8	11	9	US-09-915-306-4	Sequence 4, Appli
28	59	2.8	11	9	US-09-915-306-14	Sequence 14, Appl
29	59	2.8	11	9	US-09-915-374-4	Sequence 4, Appli
30	59	2.8	11	9	US-09-915-374-14	Sequence 14, Appl
31	59	2.8	12	10	US-09-988-493-52	Sequence 52, Appl
32	58	2.7	12	10	US-09-988-493-56	Sequence 56, Appl
33	58	2.7	12	17	US-10-700-330-193	Sequence 193, App
34	58	2.7	25	14	US-10-352-786-2311	Sequence 2311, Ap
35	58	2.7	26	14	US-10-352-786-2471	Sequence 2471, Ap
36	58	2.7	26	14	US-10-352-786-2479	Sequence 2479, Ap
37	58	2.7	27	14	US-10-352-786-2639	Sequence 2639, Ap
38	58	2.7	27	14	US-10-352-786-2647	Sequence 2647, Ap
39	58	2.7	27	14	US-10-352-786-2659	Sequence 2659, Ap
40	58	2.7	28	14	US-10-352-786-2815	Sequence 2815, Ap
41	58	2.7	28	14	US-10-352-786-2823	Sequence 2823, Ap
42	58	2.7	28	14	US-10-352-786-2835	Sequence 2835, Ap
43	58	2.7	29	14	US-10-352-786-2999	Sequence 2999, Ap
44	58	2.7	29	14	US-10-352-786-3007	Sequence 3007, Ap
45	58	2.7	29	14	US-10-352-786-3019	Sequence 3019, Ap

ALIGNMENTS

RESULT 1
US-09-915-306-2
; Sequence 2, Application US/09915306
; Patent No. US20020081310A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,306
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-306-2

Query Match 4.9%; Score 103; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 385 EDGDFNLGDLDSNSMOT 404
|||||

Db 1 EDGEDFNLDSSNSMOT 20

RESULT 2

US-09-915-374-2
; Sequence 2, Application US/09915374
; Publication No. US20020197657A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,374
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-374-2

Query Match

Best Local Similarity 4.9%; Score 103; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 EDGEDFNLDSSNSMOT 404

Db 1 EDGEDFNLDSSNSMOT 20

RESULT 3

US-09-915-306-3
; Sequence 3, Application US/09915306
; Patent No. US20020081310A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,306
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-306-3

Query Match

4.3%; Score 92; DB 9; Length 18;

Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNLDSSNS 401

Db 1 LEDGEDFNLDSSNS 18

RESULT 4

US-09-915-374-3
; Sequence 3, Application US/09915374
; Publication No. US20020197657A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,374
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-374-3

Query Match

Best Local Similarity 4.3%; Score 92; DB 9; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNLDSSNS 401

Db 1 LEDGEDFNLDSSNS 18

RESULT 5

US-09-915-306-11
; Sequence 11, Application US/09915306
; Patent No. US20020081310A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,306
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-306-11

Query Match 3.7%; Score 78; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNGLDGLDS 398
Db 1 LEDGEDFNGLDGLDS 15

RESULT 6

US-09-915-374-11
; Sequence 11, Application US/09915374
; Publication No. US20020197657A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,374
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-374-11

Query Match 3.7%; Score 78; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNGLDGLDS 398
Db 1 LEDGEDFNGLDGLDS 15

RESULT 7

US-09-915-306-13
; Sequence 13, Application US/09915306
; Patent No. US20020081310A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,306
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-306-13

Query Match 3.3%; Score 70; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 EDGEDFNGLDGLD 397
Db 1 EDGEDFNGLDGLD 13

RESULT 8

US-09-915-374-13
; Sequence 13, Application US/09915374
; Publication No. US20020197657A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,374
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-374-13

Query Match 3.3%; Score 70; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 EDGEDFNGLDGLD 397
Db 1 EDGEDFNGLDGLD 13

RESULT 9

US-09-915-306-10
; Sequence 10, Application US/09915306
; Patent No. US20020081310A1
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,306
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366

/ PRIOR FILING DATE: 1998-09-29
/ PRIOR APPLICATION NUMBER: SE 9703456-3
/ PRIOR FILING DATE: 1997-09-30
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 10
/ LENGTH: 13
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: epitope from
/ OTHER INFORMATION: mammalian protein
US-09-915-306-10

Query Match 3.2%; Score 68; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNLDAL 396
Db 1 LEDGEDFNLDAL 13

RESULT 10
US-09-915-374-10
/ Sequence 10, Application US/09915374
/ Publication No. US20020197657A1
/ GENERAL INFORMATION:
/ APPLICANT: Bjorklund, Viveka
/ APPLICANT: Bjorklund, Bertil
/ APPLICANT: Bjorklund, Peter
/ APPLICANT: Nap, Marius
/ APPLICANT: Ramaekers, Frans C.S.
/ APPLICANT: Schutte, Bert
/ TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
/ FILE REFERENCE: Bjorklund 09/162366
/ CURRENT APPLICATION NUMBER: US/09/915,374
/ CURRENT FILING DATE: 2001-07-27
/ PRIOR APPLICATION NUMBER: 09/162,366
/ PRIOR FILING DATE: 1998-07-27
/ PRIOR APPLICATION NUMBER: SE 9703456-3
/ PRIOR FILING DATE: 1997-09-30
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 10
/ LENGTH: 13
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: epitope from
/ OTHER INFORMATION: mammalian protein
US-09-915-374-10

Query Match 3.2%; Score 68; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNLDAL 396
Db 1 LEDGEDFNLDAL 13

RESULT 11
US-09-915-306-12
/ Sequence 12, Application US/09915306
/ Patent No. US20020081310A1
/ GENERAL INFORMATION:
/ APPLICANT: Bjorklund, Viveka
/ APPLICANT: Bjorklund, Bertil
/ APPLICANT: Bjorklund, Peter
/ APPLICANT: Nap, Marius
/ APPLICANT: Ramaekers, Frans C.S.
/ APPLICANT: Schutte, Bert

/ TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
/ FILE REFERENCE: Bjorklund 09/162366
/ CURRENT APPLICATION NUMBER: US/09/915,306
/ CURRENT FILING DATE: 2001-07-27
/ PRIOR APPLICATION NUMBER: 09/162,366
/ PRIOR FILING DATE: 1998-09-29
/ PRIOR APPLICATION NUMBER: SE 9703456-3
/ PRIOR FILING DATE: 1997-09-30
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 12
/ LENGTH: 12
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: epitope from
/ OTHER INFORMATION: mammalian protein
US-09-915-306-12

Query Match 3.1%; Score 65; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 DGEDFNLDALD 397
Db 1 DGEDFNLDALD 12

RESULT 12
US-09-915-374-12
/ Sequence 12, Application US/09915374
/ Publication No. US20020197657A1
/ GENERAL INFORMATION:
/ APPLICANT: Bjorklund, Viveka
/ APPLICANT: Bjorklund, Bertil
/ APPLICANT: Bjorklund, Peter
/ APPLICANT: Nap, Marius
/ APPLICANT: Ramaekers, Frans C.S.
/ APPLICANT: Schutte, Bert
/ TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
/ FILE REFERENCE: Bjorklund 09/162366
/ CURRENT APPLICATION NUMBER: US/09/915,374
/ CURRENT FILING DATE: 2001-07-27
/ PRIOR APPLICATION NUMBER: 09/162,366
/ PRIOR FILING DATE: 1998-07-27
/ PRIOR APPLICATION NUMBER: SE 9703456-3
/ PRIOR FILING DATE: 1997-09-30
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 12
/ LENGTH: 12
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: epitope from
/ OTHER INFORMATION: mammalian protein
US-09-915-374-12

Query Match 3.1%; Score 65; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 DGEDFNLDALD 397
Db 1 DGEDFNLDALD 12

RESULT 13
US-09-988-493-57
/ Sequence 57, Application US/09988493
/ Publication No. US20030064419A1
/ GENERAL INFORMATION:
/ APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri

; APPLICANT: O'Hare, Michael John
; APPLICANT: Page, Martin John
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Waterfield, Michael Derek
; TITLE OF INVENTION: Proteins, Genes, and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Breast Cancer
; FILE REFERENCE: 2543-1-024
; CURRENT APPLICATION NUMBER: US/09/988,493
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: PCT/GB01/01219
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: GB 0006695.1
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: GB 0007265.2
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 13
; TYPE: PRT
; ORGANISM: homo sapien
US-09-988-493-57

Query Match 3.0%; Score 64; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SLGSVQAPSYGAR 27
Db 1 SLGSVQAPSYGAR 13

RESULT 14

US-10-700-330-215
; Sequence 215, Application US/10700330
; Publication No. US20040203022A1

GENERAL INFORMATION:

; APPLICANT: Herath, Mudiyansele Athula Chandrasiri Herath
; APPLICANT: Page, Martin John
; TITLE OF INVENTION: Proteins and Genes For Diagnosis And Treatment of ErbB2-Related C
; FILE REFERENCE: 2543-1-031
; CURRENT APPLICATION NUMBER: US/10/700,330
; CURRENT FILING DATE: 2003-11-03
; PRIOR APPLICATION NUMBER: GB 0110886.9
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: GB 0128183.1
; PRIOR FILING DATE: 2001-11-23
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 215
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-700-330-215

Query Match 3.0%; Score 64; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SLGSVQAPSYGAR 27
Db 1 SLGSVQAPSYGAR 13

RESULT 15

US-09-988-493-59
; Sequence 59, Application US/09988493
; Publication No. US20030064419A1

GENERAL INFORMATION:

; APPLICANT: Herath, Mudiyansele Athula Chandrasiri
; APPLICANT: O'Hare, Michael John
; APPLICANT: Page, Martin John
; APPLICANT: Parekh, Rajesh Bhikhu

; APPLICANT: Waterfield, Michael Derek
; TITLE OF INVENTION: Proteins, Genes, and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Breast Cancer
; FILE REFERENCE: 2543-1-024
; CURRENT APPLICATION NUMBER: US/09/988,493
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: PCT/GB01/01219
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: GB 0006695.1
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: GB 0007265.2
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 13
; TYPE: PRT
; ORGANISM: homo sapien
US-09-988-493-59

Query Match 2.9%; Score 61; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 TVQSLEIDLDSMR 314
Db 1 TVQSLEIDLDSMR 13

Search completed: December 14, 2004, 09:36:53
Job time : 144 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 09:22:06 ; Search time 39 Seconds
(without alignments)
731.198 Million cell updates/sec

Title: US-10-026-001-1
Perfect score: 2120
Sequence: 1 MSFTTRSTSTNYRSLGSVQ.....RRVDGKVSEVNDTKVLRH 430

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 223511

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	5.0	22	1 US-07-934-656A-7	Sequence 7, Appli
2	103	4.9	20	3 US-09-162-366C-2	Sequence 2, Appli
3	103	4.9	20	4 US-09-915-374-2	Sequence 2, Appli
4	103	4.9	20	4 US-09-915-306-2	Sequence 2, Appli
5	93	4.4	22	1 US-07-934-656A-22	Sequence 22, Appli
6	92	4.3	18	3 US-09-162-366C-3	Sequence 3, Appli
7	92	4.3	18	4 US-09-915-374-3	Sequence 3, Appli
8	92	4.3	18	4 US-09-915-306-3	Sequence 3, Appli
9	86	4.1	22	1 US-07-934-656A-6	Sequence 6, Appli
10	85	4.0	22	1 US-07-934-656A-3	Sequence 3, Appli
11	85	4.0	22	1 US-07-934-656A-4	Sequence 4, Appli
12	82	3.9	22	1 US-07-934-656A-2	Sequence 2, Appli
13	81	3.8	22	1 US-07-934-656A-5	Sequence 5, Appli
14	78	3.7	15	3 US-09-162-366C-11	Sequence 11, Appli
15	78	3.7	15	4 US-09-915-374-11	Sequence 11, Appli
16	78	3.7	15	4 US-09-915-306-11	Sequence 11, Appli
17	70.5	3.3	21	1 US-07-934-656A-30	Sequence 30, Appli
18	70	3.3	13	3 US-09-162-366C-13	Sequence 13, Appli
19	70	3.3	13	4 US-09-915-374-13	Sequence 13, Appli
20	70	3.3	13	4 US-09-915-306-13	Sequence 13, Appli
21	68	3.2	13	3 US-09-162-366C-10	Sequence 10, Appli
22	68	3.2	13	4 US-09-915-374-10	Sequence 10, Appli
23	68	3.2	13	4 US-09-915-306-10	Sequence 10, Appli
24	65	3.1	12	3 US-09-162-366C-12	Sequence 12, Appli
25	65	3.1	12	4 US-09-915-374-12	Sequence 12, Appli
26	65	3.1	12	4 US-09-915-306-12	Sequence 12, Appli
27	61	2.9	16	4 US-09-809-003A-14	Sequence 14, Appli

28	59	2.8	11	3 US-09-162-366C-4	Sequence 4, Appli
29	59	2.8	11	3 US-09-162-366C-14	Sequence 14, Appli
30	59	2.8	11	4 US-09-915-374-4	Sequence 4, Appli
31	59	2.8	11	4 US-09-915-374-14	Sequence 14, Appli
32	59	2.8	11	4 US-09-915-306-4	Sequence 4, Appli
33	59	2.8	11	4 US-09-915-306-14	Sequence 14, Appli
34	58.5	2.8	30	1 US-08-425-069-48	Sequence 48, Appli
35	58.5	2.8	30	1 US-08-425-069-60	Sequence 60, Appli
36	58.5	2.8	30	2 US-08-317-844B-48	Sequence 48, Appli
37	58.5	2.8	30	2 US-08-317-844B-60	Sequence 60, Appli
38	58	2.7	22	1 US-07-934-656A-19	Sequence 19, Appli
39	56.5	2.7	28	1 US-08-425-069-50	Sequence 50, Appli
40	56.5	2.7	28	2 US-08-317-844B-50	Sequence 50, Appli
41	55.5	2.6	30	1 US-08-425-069-53	Sequence 53, Appli
42	55.5	2.6	30	2 US-08-317-844B-53	Sequence 53, Appli
43	55	2.6	16	1 US-07-934-656A-1	Sequence 1, Appli
44	55	2.6	22	4 US-08-406-824A-24	Sequence 24, Appli
45	54.5	2.6	30	1 US-08-425-069-61	Sequence 61, Appli

ALIGNMENTS

RESULT 1
US-07-934-656A-7
; Sequence 7, Application US/07934656A
; Patent No. 5500347
; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; APPLICANT: FRANK, Werner W.
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; TITLE OF INVENTION: CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,656A
; FILING DATE: 27-JAN-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 23 945.4
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-3003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-7

Query Match 5.0%; Score 105; DB 1; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.069;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 77 IQNEKETMQSLNDRLASLYLDRV 98

Db 1 IQNEKETMQSLNDRSLASYLDKV 22
|||||

RESULT 2
US-09-162-366C-2
; Sequence 2, Application US/09162366C
; Patent No. 6296850
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/162,366C
; CURRENT FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/060,556
; PRIOR FILING DATE: 1997-09-30
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-162-366C-2

Query Match 4.9%; Score 103; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 EDGEDFNGLGDALDSSNSMQT 404
|||||
Db 1 EDGEDFNGLGDALDSSNSMQT 20

RESULT 3
US-09-915-374-2
; Sequence 2, Application US/09915374
; Patent No. 6706488
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,374
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-374-2

Query Match 4.9%; Score 103; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 EDGEDFNGLGDALDSSNSMQT 404
|||||
Db 1 EDGEDFNGLGDALDSSNSMQT 20

RESULT 4
US-09-915-306-2
; Sequence 2, Application US/09915306
; Patent No. 6716968
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,306
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-306-2

Query Match 4.9%; Score 103; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 EDGEDFNGLGDALDSSNSMQT 404
|||||
Db 1 EDGEDFNGLGDALDSSNSMQT 20

RESULT 5
US-07-934-656A-22
; Sequence 22, Application US/07934656A
; Patent No. 5500347
; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; APPLICANT: FRANK, Werner W.
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; TITLE OF INVENTION: CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,656A

;; FILING DATE: 27-JAN-1993
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: DE P 40 23 945.4
;; FILING DATE: 27-JUL-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murray, Robert B.
;; REGISTRATION NUMBER: 22,980
;; REFERENCE/DOCKET NUMBER: P564-3003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)638-5000
;; TELEFAX: (202)638-4810
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-07-934-656A-22

Query Match 4.4%; Score 93; DB 1; Length 22;
Best Local Similarity 90.9%; Pred. No. 0.58;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 233 LTVVDAPKSQLAKIMADIRA 254
Db 1 LTVVDAPKSQLSIIMADIRA 22

RESULT 6
US-09-162-366C-3
; Sequence 3, Application US/09162366C
; Patent No. 6296850
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/162,366C
; CURRENT FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/060,556
; PRIOR FILING DATE: 1997-09-30
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-162-366C-3

Query Match 4.3%; Score 92; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNGLDSSNS 401
Db 1 LEDGEDFNGLDSSNS 18

RESULT 7
US-09-915-374-3
; Sequence 3, Application US/09915374
; Patent No. 6706488
; GENERAL INFORMATION:

;; APPLICANT: Bjorklund, Viveka
;; APPLICANT: Bjorklund, Bertil
;; APPLICANT: Bjorklund, Peter
;; APPLICANT: Nap, Marius
;; APPLICANT: Ramaekers, Frans C.S.
;; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,374
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-374-3

Query Match 4.3%; Score 92; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNGLDSSNS 401
Db 1 LEDGEDFNGLDSSNS 18

RESULT 8
US-09-915-306-3
; Sequence 3, Application US/09915306
; Patent No. 6716968
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915,306
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-306-3

Query Match 4.3%; Score 92; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNGLDSSNS 401
Db 1 LEDGEDFNGLDSSNS 18

RESULT 9
US-07-934-656A-6
; Sequence 6, Application US/07934656A
; Patent No. 5500347
; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; APPLICANT: FRANK, Werner W.
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; TITLE OF INVENTION: CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,656A
; FILING DATE: 27-JAN-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 23 945.4
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-3003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-6
Query Match 4.1%; Score 86; DB 1; Length 22;
Best Local Similarity 85.0%; Pred. No. 2;
Matches 17; Conservative 2; Mismatches 1; Indels 1; Gaps 0;
OY 79 NEKETMQSLNDRSLASYLDKV 98
Db 3 NEKVTMQNLNDRSLASYLDKV 22
RESULT 10
US-07-934-656A-3
; Sequence 3, Application US/07934656A
; Patent No. 5500347
; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; APPLICANT: FRANK, Werner W.
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; TITLE OF INVENTION: CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,656A
; FILING DATE: 27-JAN-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 23 945.4
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-3003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-6
Query Match 4.1%; Score 86; DB 1; Length 22;
Best Local Similarity 85.0%; Pred. No. 2;
Matches 17; Conservative 2; Mismatches 1; Indels 1; Gaps 0;
OY 79 NEKETMQSLNDRSLASYLDKV 98
Db 3 NEKVTMQNLNDRSLASYLDKV 22
RESULT 11
US-07-934-656A-4
; Sequence 4, Application US/07934656A
; Patent No. 5500347
; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; APPLICANT: FRANK, Werner W.
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; TITLE OF INVENTION: CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,656A
; FILING DATE: 27-JAN-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 23 945.4
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-3003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-3

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/934,656A
FILING DATE: 27-JAN-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 40 23 945.4
FILING DATE: 27-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Robert B.
REGISTRATION NUMBER: 22,980
REFERENCE/DOCKET NUMBER: P564-3003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-934-656A-3
Query Match 4.0%; Score 85; DB 1; Length 22;
Best Local Similarity 85.0%; Pred. No. 2.4;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 79 NEKETMQSLNDRSLASYLDKV 98
Db 3 NEKVTMQNLNDRSLASYLDKV 22
RESULT 11
US-07-934-656A-4
; Sequence 4, Application US/07934656A
; Patent No. 5500347
; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; APPLICANT: FRANK, Werner W.
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; TITLE OF INVENTION: CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,656A
; FILING DATE: 27-JAN-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 23 945.4
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-3003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-3

; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-4

Query Match 4.0%; Score 85; DB 1; Length 22;
Best Local Similarity 85.0%; Pred. No. 2.4;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 79 NEKETMQSLNDRSLASVLDV 98
||| |||:|||||||:|
Db 3 NEKLTQNLNDRSLASVLDKV 22

RESULT 12

US-07-934-656A-2
; Sequence 2, Application US/07934656A
; Patent No. 5500347

; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07934,656A
; FILING DATE: 27-JAN-1993

; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 23 945.4
; FILING DATE: 27-JUL-1990

; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-3003
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-2

Query Match 3.9%; Score 82; DB 1; Length 22;
Best Local Similarity 72.7%; Pred. No. 4.1;
Matches 16; Conservative 4; Mismatches 2; Indels 2; Gaps 0;

QY 77 IQNEKETMQSLNDRSLASVLDV 98
: ||| |||:|||||||:|
Db 1 VGSEKVTMQNLNDRSLASVLDKV 22

RESULT 13

US-07-934-656A-5
; Sequence 5, Application US/07934656A

; Patent No. 5500347
; GENERAL INFORMATION:
; APPLICANT: MOLL, Roland
; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF
; CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF
; ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07934,656A
; FILING DATE: 27-JAN-1993

; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 23 945.4
; FILING DATE: 27-JUL-1990

; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-3003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-934-656A-5

Query Match 3.8%; Score 81; DB 1; Length 22;
Best Local Similarity 80.0%; Pred. No. 4.9;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 79 NEKETMQSLNDRSLASVLDV 98
||| |||:|||||||:|
Db 3 NEKITQNLNDRSLASVLEKV 22

RESULT 14

US-09-162-366C-11
; Sequence 11, Application US/09162366C
; Patent No. 6296850

; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius

; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/162,366C

; CURRENT FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/060,556
; PRIOR FILING DATE: 1997-09-30

; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11

;
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-162-366C-11

Query Match 3.7%; Score 78; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNLDALDS 398
|||
Db 1 LEDGEDFNLDALDS 15

RESULT 15
US-09-915-374-11
; Sequence 11, Application US/09915374
; Patent No. 6706488
; GENERAL INFORMATION:
; APPLICANT: Bjorklund, Viveka
; APPLICANT: Bjorklund, Bertil
; APPLICANT: Bjorklund, Peter
; APPLICANT: Nap, Marius
; APPLICANT: Ramaekers, Frans C.S.
; APPLICANT: Schutte, Bert
; TITLE OF INVENTION: Apoptosis-Related Compounds and Their Use
; FILE REFERENCE: Bjorklund 09/162366
; CURRENT APPLICATION NUMBER: US/09/915.374
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 09/162,366
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: SE 9703456-3
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: epitope from
; OTHER INFORMATION: mammalian protein
US-09-915-374-11

Query Match 3.7%; Score 78; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LEDGEDFNLDALDS 398
|||
Db 1 LEDGEDFNLDALDS 15

Search completed: December 14, 2004, 09:34:24
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: December 14, 2004, 09:13:15 ; Search time 151 Seconds
(without alignments)
1021.547 Million cell updates/sec

Title: US-10-026-001-1
Perfect score: 2120
Sequence: 1 MSFTTRSTPSTNYRSLGSVQ.....RRVDGKVVSEINPTKVLRH 430

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 777192

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04: *
1: Geneseq1980s: *
2: Geneseq1990s: *
3: Geneseq2000s: *
4: Geneseq2001s: *
5: Geneseq2002s: *
6: Geneseq2003as: *
7: Geneseq2003bs: *
8: Geneseq2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	139	6.6	29	2	AAR85703	Human cyt
2	131	6.2	28	4	ABG18695	Novel hum
3	125	5.9	25	2	AAR85704	Human cyt
4	119	5.6	25	2	AAR85706	Human cyt
5	114	5.4	24	2	AAR85705	Human cyt
6	103	4.9	20	2	AAR85707	Cytokerat
7	103	4.9	21	2	AAR85707	Human cyt
8	100	4.7	20	6	AAE38127	Human cyt
9	92	4.3	18	2	AAW93848	Cytokerat
10	76	3.6	15	6	AAE38126	Human cyt
11	66	3.1	27	7	ADJ81155	Self-asse
12	66	3.1	27	7	ADJ81362	Self-asse
13	66	3.1	27	7	ADJ81232	Self-asse
14	66	3.1	27	7	ADJ81293	Self-asse
15	64	3.0	13	4	AAU68081	Human Bre
16	64	3.0	13	6	ABP99026	ErbB2 cel
17	61	2.9	13	4	AAU68083	Human Bre
18	61	2.9	13	6	ABP99024	ErbB2 cel
19	61	2.9	16	6	AAO26394	Psoriasis
20	61	2.9	27	7	ADF21417	Bioactive
21	61	2.9	28	7	ADF21593	Bioactive
22	61	2.9	29	7	ADF21777	Bioactive
23	61	2.9	30	7	ADF22198	Bioactive
24	60.5	2.9	30	4	AAG77868	MHC class
25	60.5	2.9	30	6	AAE37195	Linker pe

26	60.5	2.9	30	8	ADN00801	Epitope p
27	60.5	2.9	30	8	ADO21853	Histone a
28	60.5	2.9	30	8	ADP48740	Linker pe
29	60	2.8	27	7	ADF21415	Bioactive
30	60	2.8	28	7	ADF21591	Bioactive
31	60	2.8	29	7	ADF21775	Bioactive
32	60	2.8	30	7	ADF22196	Bioactive
33	59	2.8	11	2	AAW93849	Cytokerat
34	59	2.8	12	4	AAU68074	Human Bre
35	59	2.8	12	4	AAU68076	Human Bre
36	58.5	2.8	30	3	AAU59111	N. clavi
37	58.5	2.8	30	3	AAU59123	N. clavi
38	58	2.7	12	4	AAU68079	Human Bre
39	58	2.7	12	6	ABP99025	ErbB2 cel
40	58	2.7	25	7	ADF21073	Bioactive
41	58	2.7	26	7	ADF21233	Bioactive
42	58	2.7	26	7	ADF21241	Bioactive
43	58	2.7	27	7	ADF21401	Bioactive
44	58	2.7	27	7	ADF21409	Bioactive
45	58	2.7	27	7	ADF21421	Bioactive

ALIGNMENTS

RESULT 1
AAR85703
ID AAR85703 standard; peptide; 29 AA.

XX	AC	AAR85703;
XX	DT	21-JUN-1996 (first entry)
XX	DE	Human cytokeratin 18 fragment.
XX	KW	Human; cytokeratin 18; indication; tumour; carcinoma; liver; kidney;
KW	KW	disease; proliferation; repair processes; diagnosis;
KW	KW	monitoring treatment; antibodies; immunoassay reagents; in vivo;
XX	OS	Homo sapiens.
XX	PN	WO9531728-A1.
XX	PD	23-NOV-1995.
XX	PF	15-MAY-1995; 95WO-SE000532.
XX	PR	17-MAY-1994; 94SE-00001687.
XX	PA	04-JAN-1995; 95SE-00000023.
XX	PI	(BEKI-) BEKI AB.
XX	DR	Bjoerklund P, Rydlander L;
XX	PT	WPI; 1996-011077/01.
XX	PS	Diagnosing cancer from elevated levels of cytokeratin fragments in body
XX	CC	fluid - also new fragments and their fusion proteins; nucleic acid
XX	CC	encoding them, antibodies, etc.
XX	CC	Claim 4; Page 26; 31pp; English.
XX	CC	Elevated levels of AAR85703/04 (human cytokeratin 18 fragments) in a body
XX	CC	fluid may indicate tumour (esp. carcinoma) cell activity, but may also be
XX	CC	encountered in cases of liver or kidney disease, or of proliferative
XX	CC	activity associated with repair processes. The peptides can therefore be
XX	CC	used for diagnosis, monitoring treatment and to generate antibodies,
XX	CC	which can be used as immunoassay reagents, or for in vivo tumour
XX	CC	localisation
XX	SQ	Sequence 29 AA;

Query Match
Best Local Similarity 100.0%; Pred. No. 0.0057;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 ASLENSLREVEARYALQMEQLNGILLHLE 346
Db 1 ASLENSLREVEARYALQMEQLNGILLHLE 29

RESULT 2
ABG18695
ID ABG18695 standard; protein; 28 AA.
XX
AC ABG18695;
DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #18686.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS82882.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.
XX
PS Claim 20; SEQ ID NO 49054; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 28 AA;

Query Match
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IFANTVDNARIVLQIDNARLAADPRV 166
Db 1 IFANTVDNARIVLQIDNARLAADPRV 27

RESULT 3
AAR85704
ID AAR85704 standard; peptide; 25 AA.
XX
AC AAR85704;
DT 21-JUN-1996 (first entry)
DE Human cytokeratin 18 fragment.
XX
KW Human; cytokeratin 18; indication; tumour; carcinoma; liver; kidney; disease; proliferation; repair processes; diagnosis; monitoring treatment; antibodies; immunoassay reagents; in vivo; localisation.
XX
OS Homo sapiens.
XX
PN WO9531728-A1.
XX
PD 23-NOV-1995.
XX
PF 15-MAY-1995; 95WO-SE000532.
XX
PR 17-MAY-1994; 94SE-00001687.
XX
PR 04-JAN-1995; 95SE-00000023.
XX
PA (BEKI-) BEKI AB.
XX
PI Bjoerklund P, Rydlander L;
XX
DR WPI; 1996-011077/01.
XX
PT Diagnosing cancer from elevated levels of cytokeratin fragments in body fluid - also new fragments and their fusion proteins, nucleic acid encoding them, antibodies, etc.
XX
PS Claim 4; Page 26; 31pp; English.
XX
CC Elevated levels of AAR85703/04 (human cytokeratin 18 fragments) in a body fluid may indicate tumour (esp. carcinoma) cell activity, but may also be encountered in cases of liver or kidney disease, or of proliferative activity associated with repair processes. The peptides can therefore be used for diagnosis, monitoring treatment and to generate antibodies, which can be used as immunoassay reagents, or for in vivo tumour localisation
XX
SQ Sequence 25 AA;

Query Match
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 QKTTTRRIVDGKVVSETNDTKVLRH 430
Db 1 QKTTTRRIVDGKVVSETNDTKVLRH 25

RESULT 4
AAR85706
ID AAR85706 standard; peptide; 25 AA.
XX
AC AAR85706;
DT 21-JUN-1996 (first entry)
DE Human cytokeratin 18 fragment homologue.

XX Human; cytokeratin 18; indication; tumour; carcinoma; liver; kidney;
KW disease; proliferation; repair processes; diagnosis;
KW monitoring treatment; antibodies; immunoassay reagents; in vivo;
KW localisation; homologue.
XX
OS Homo sapiens.
XX WO9531728-A1.
XX 23-NOV-1995.
XX
PF 15-MAY-1995; 95WO-SE000532.
XX
PR 17-MAY-1994; 94SE-00001687.
PR 04-JAN-1995; 95SE-00000023.
XX
PA (BEKI-) BEKI AB.
XX Bjoerklund P, Rydlander L;
XX WPI; 1996-011077/01.
DR
XX
PT Diagnosing cancer from elevated levels of cytokeratin fragments in body
PT fluid - also new fragments and their fusion proteins, nucleic acid
PT encoding them, antibodies, etc.
XX
PS Disclosure; Page 6; 31pp; English.
XX
CC Elevated levels of AAR85703/04 (human cytokeratin 18 fragments, or their
CC homologues AAR85705-7) in a body fluid may indicate tumour (esp.
CC carcinoma) cell activity, but may also be encountered in cases of liver
CC or kidney disease, or of proliferative activity associated with repair
CC processes. The peptides can therefore be used for diagnosis, monitoring
CC treatment and to generate antibodies, which can be used as immunoassay
CC reagents, or for in vivo tumour localisation
XX
SQ Sequence 25 AA;

Query Match 5.6%; Score 119; DB 2; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.12;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 322 NSLREVEARYALQMEQLNGILLHLE 346
DB 1 NSLREVEARYALQMEQLNGILLHLE 25

RESULT 5
AAR85705
ID AAR85705 standard; peptide; 24 AA.
XX
AC AAR85705;
XX
DT 21-JUN-1996 (first entry)
XX
DE Human cytokeratin 18 fragment homologue.
XX
KW Human; cytokeratin 18; indication; tumour; carcinoma; liver; kidney;
KW disease; proliferation; repair processes; diagnosis;
KW monitoring treatment; antibodies; immunoassay reagents; in vivo;
KW localisation; homologue.
XX
OS Homo sapiens.
XX
XX WO9531728-A1.
XX
PD 23-NOV-1995.
XX
PF 15-MAY-1995; 95WO-SE000532.
XX
PR 17-MAY-1994; 94SE-00001687.
PR 04-JAN-1995; 95SE-00000023.

XX (BEKI-) BEKI AB.
XX Bjoerklund P, Rydlander L;
XX WPI; 1996-011077/01.
DR
XX
PT Diagnosing cancer from elevated levels of cytokeratin fragments in body
PT fluid - also new fragments and their fusion proteins, nucleic acid
PT encoding them, antibodies, etc.
XX
PS Disclosure; Page 6; 31pp; English.
XX
CC Elevated levels of AAR85703/04 (human cytokeratin 18 fragments, or their
CC homologues AAR85705-7) in a body fluid may indicate tumour (esp.
CC carcinoma) cell activity, but may also be encountered in cases of liver
CC or kidney disease, or of proliferative activity associated with repair
CC processes. The peptides can therefore be used for diagnosis, monitoring
CC treatment and to generate antibodies, which can be used as immunoassay
CC reagents, or for in vivo tumour localisation
XX
SQ Sequence 24 AA;

Query Match 5.4%; Score 114; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 ASLENSLREVEARYALQMEQLNGI 341
DB 1 ASLENSLREVEARYALQMEQLNGI 24

RESULT 6
AAW93847
ID AAW93847 standard; peptide; 20 AA.
XX
AC AAW93847;
XX
DT 25-JUN-1999 (first entry)
XX
DE Cytokeratin 18 apoptosis-related antigenic oligopeptide.
XX
KW Monoclonal antibody; Mab; M30; human; apoptosis-related; antigenic;
KW cytokeratin 18; CK18; apoptosis inhibition; diagnosis; cancer; therapy;
KW degenerative disease; treatment; anorexia; AIDS; Alzheimer's disease;
KW organ transplantation; psoriasis; cytostatic; anti-HIV; neuroprotective;
KW anti-psoriatic; immunosuppressant.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916789-A1.
XX
PD 08-APR-1999.
XX
PF 25-SEP-1998; 98WO-SE001721.
XX
PR 30-SEP-1997; 97SE-00003546.
PR 30-SEP-1997; 97US-0060556P.
XX
PA (BEKI-) BEKI PUBL AB.
XX
PI Bjoerklund V, Bjoerklund B, Bjoerklund P, Nap M, Ramaekers FCS;
PI Schutte B;
XX
DR WPI; 1999-277070/23.
XX
PT Apoptosis-related antigenic compounds, derived from cytokeratin 18.
XX
PS Disclosure; Page 21; 32pp; English.
XX
CC This invention describes apoptosis-related antigenic compounds which are
CC fragments of human cytokeratin 18 and which can be used for inhibition of

CC cell apoptosis. Antibodies recognizing the antigenic peptide are useful
 CC for the stimulation of cell apoptosis. The rate of cell apoptosis
 CC determined in the assay is useful in diagnosis of degenerative diseases
 CC and cancer and/or monitoring the effect of therapy. The antibodies and
 CC antigenic peptides are also used to treat cancer and degenerative
 CC diseases involving apoptosis, e.g. anorexia, AIDS, transplantation of
 CC organs, psoriasis and Alzheimer's disease. The monoclonal antibody M30
 CC specifically recognizes apoptotic cells in the early stages of the
 CC process and is applicable to fresh and also formalin fixed, paraffin
 CC embedded tissue sections of routinely obtained biopsies. The products of
 CC the invention have antigenic, cytostatic, anti-HIV, neuroprotective, anti
 CC -psoriatic and immunosuppressant activity

XX Sequence: 20 AA;

Query Match 4.9%; Score 103; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 385 EDGEDFNLDSSNSMQT 404
 Db 1 EDGEDFNLDSSNSMQT 20

RESULT 7

AAR85707

ID AAR85707 standard; peptide; 21 AA.

XX AAR85707;

AC AAR85707;

XX 21-JUN-1996 (first entry)

XX Human cytokeratin 18 fragment homologue.

XX Human; cytokeratin 18; indication; tumour; carcinoma; liver; kidney;
 KW disease; proliferation; repair processes; diagnosis;
 KW monitoring treatment; antibodies; immunoassay reagents; in vivo;
 KW localisation; homologue.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "opt. Lys, Gly, Ala or Glu"
 FT Misc-difference 2 /note= "opt. Lys or Gly, pref. Ala or Glu"
 FT Misc-difference 3 /note= "opt. Lys, Gly, Ala or Tyr"
 FT Misc-difference 4 /note= "opt. Lys or Gly, pref. Ala or Glu"
 FT Misc-difference 5 /note= "opt. Lys, Gly, Ala or Tyr"
 FT Misc-difference 6 /note= "opt. Lys, Gly or Tyr, pref. Glu"
 FT Misc-difference 7 /note= "opt. Glu or Gly, pref. Ala or Lys"
 FT Misc-difference 8 /note= "opt. Lys, Gly, Ala or Glu"
 FT Misc-difference 9 /note= "opt. Lys, Gly or Tyr, pref. Glu"
 FT Misc-difference 10 /note= "opt. Lys, pref. Ala or Glu"
 FT Misc-difference 11 /note= "opt. Lys, Ala or Glu"
 FT Misc-difference 12 /note= "opt. Glu or Gly, pref. Ala or Lys"
 FT Misc-difference 13 /note= "opt. Lys, Ala or Glu"
 FT Misc-difference 14 /note= "opt. Tyr, Ala or Glu"
 FT Misc-difference 15 /note= "opt. Lys, Gly, Ala or Glu"
 FT Misc-difference 16 /note= "opt. Glu or Gly, pref. Ala or Lys"
 FT Misc-difference 17 /note= "opt. Tyr, Ala or Glu"
 FT Misc-difference 18 /note= "opt. Lys, Gly, Ala or Glu"
 FT Misc-difference 19 /note= "opt. Glu or Gly, pref. Ala or Lys"
 FT Misc-difference 20 /note= "opt. Lys, Gly, Ala or Glu"
 FT Misc-difference 21 /note= "opt. Glu or Gly, pref. Ala or Lys"

FT /note= "opt. Glu or Lys, pref. Gly or Ala"

XX WO9531728-A1.

XX 23-NOV-1995.

XX 15-MAY-1995; 95WO-SE0000532.

XX 17-MAY-1994; 94SE-00001687.

XX 04-JAN-1995; 95SE-00000023.

XX (BEKI-) BEKI AB.

XX Bjoerklund P, Rydlander L;

XX WPI; 1996-011077/01.

XX Diagnosing cancer from elevated levels of cytokeratin fragments in body
 PT fluid - also new fragments and their fusion proteins, nucleic acid
 PT encoding them, antibodies, etc.

XX Disclosure; Page 6; 31pp; English.

XX Elevated levels of AAR85703/04 (human cytokeratin 18 fragments, or their
 CC homologues AAR85705-7) in a body fluid may indicate tumour (esp.
 CC carcinoma) cell activity, but may also be encountered in cases of liver
 CC or kidney disease, or of proliferative activity associated with repair
 CC processes. The peptides can therefore be used for diagnosis, monitoring
 CC treatment and to generate antibodies, which can be used as immunoassay
 CC reagents, or for in vivo tumour localisation

XX Sequence 21 AA;

Query Match 4.9%; Score 103; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 324 LREVEARYALQMEQLNGILLH 344

Db 1 LREVEARYALQMEQLNGILLH 21

RESULT 8

AAE38127

ID AAE38127 standard; peptide; 20 AA.

XX AAE38127;

XX 06-NOV-2003 (first entry)

XX Human cytokeratin K18 N-terminal peptide #2.

XX Human; cancer-associated epitope; cytokeratin K8; cytokeratin K18;
 KW adenocarcinoma; therapy; cancer.

XX Homo sapiens.

XX WO2003057168-A2.

XX 17-JUL-2003.

XX 03-JAN-2003; 2003WO-US0000297.

XX 03-JAN-2002; 2002US-0345208P.

XX (SCRI) SCRIPPS RES INST.

XX Ditzel H, Jensenius JC;

XX WPI; 2003-598315/56.

XX Novel isolated cancer-associated epitope comprising two separate
 PT polypeptides, a cytokeratin 8 polypeptide and a cytokeratin 18

PT polypeptide, useful as component of vaccine for preventing or treating
PT adenocarcinoma.
XX
PS Example 2; Fig 3B; 155pp; English.
XX
CC The invention provides a cancer-associated epitope comprising two
CC separate polypeptides, a cytokeratin 8 polypeptide and a cytokeratin 18
CC polypeptide. Vaccine composition of the invention is useful for treating
CC or preventing colon adenocarcinoma, ovarian adenocarcinoma, renal
CC adenocarcinoma, mammary adenocarcinoma, lung adenocarcinoma, pancreatic
CC adenocarcinoma or non-seminoma testis carcinoma. The invention is also
CC useful for preparing a medicament for treating or preventing cancer in a
CC mammal. The present sequence is human cytokeratin K18 N-terminal peptide
XX
SQ Sequence 20 AA;
Query Match 4.7%; Score 100; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 68 AGGLAGMGGIQNEKETMOSL 87
Db 1 AGGLAGMGGIQNEKETMOSL 20
RESULT 9
AAW93848
ID AAW93848 standard; peptide; 18 AA.
XX
AC AAW93848;
XX
DT 25-JUN-1999 (first entry)
XX
DE Cytokeratin 18 apoptosis-related antigenic oligopeptide 2.
XX
KW Monoclonal antibody; Mab; M30; human; apoptosis-related; antigenic;
KW cytokeratin 18; CK18; apoptosis inhibition; diagnosis; cancer; therapy;
KW degenerative disease; treatment; anorexia; AIDS; Alzheimer's disease;
KW organ transplantation; psoriasis; cytostatic; anti-HIV; neuroprotective;
KW anti-psoriatic; immunosuppressant.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916789-A1.
XX
PD 08-APR-1999.
XX
PF 25-SEP-1998; 98WO-SE001721.
XX
PR 30-SEP-1997; 97SE-00003546.
PR 30-SEP-1997; 97US-0060556P.
XX
PA (BEKI-) BEKI PUBL AB.
XX
PI Bjoerklund V, Bjoerklund B, Bjoerklund P, Nap M, Ramaekers FCS;
PI Schutte B;
XX
DR WPI; 1999-277070/23.
XX
PT Apoptosis-related antigenic compounds, derived from cytokeratin 18.
XX
PS Disclosure; Page 21; 32pp; English.
XX
CC This invention describes apoptosis-related antigenic compounds which are
CC fragments of human cytokeratin 18 and which can be used for inhibition of
CC cell apoptosis. Antibodies recognizing the antigenic peptide are useful
CC for the stimulation of cell apoptosis. The rate of cell apoptosis
CC determined in the assay is useful in diagnosis of degenerative diseases
CC and cancer and/or monitoring the effect of therapy. The antibodies and
CC antigenic peptides are also used to treat cancer and degenerative
CC diseases involving apoptosis, e.g. anorexia, AIDS, transplantation of
CC organs, psoriasis and Alzheimer's disease. The monoclonal antibody M30

CC specifically recognizes apoptotic cells in the early stages of the
CC process and is applicable to fresh and also formalin fixed, paraffin
CC embedded tissue sections of routinely obtained biopsies. The products of
CC the invention have antigenic, cytostatic, anti-HIV, neuroprotective, anti
CC -psoriatic and immunosuppressant activity
XX
SQ Sequence 18 AA;
Query Match 4.3%; Score 92; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 384 LEDGEDFNGLDSSNS 401
Db 1 LEDGEDFNGLDSSNS 18
RESULT 10
AAE38126
ID AAE38126 standard; peptide; 15 AA.
XX
AC AAE38126;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human cytokeratin K18 N-terminal peptide #1.
XX
KW Human; cancer-associated epitope; cytokeratin K8; cytokeratin K18;
KW adenocarcinoma; therapy; cancer.
XX
OS Homo sapiens.
XX
PN WO2003057168-A2.
XX
PD 17-JUL-2003.
XX
PF 03-JAN-2003; 2003WO-US000297.
XX
PR 03-JAN-2002; 2002US-0345208P.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Ditzel H, Jensenius JC;
XX
DR WPI; 2003-598315/56.
XX
PT Novel isolated cancer-associated epitope comprising two separate
PT polypeptides, a cytokeratin 8 polypeptide and a cytokeratin 18
PT polypeptide, useful as component of vaccine for preventing or treating
PT adenocarcinoma.
XX
PS Example 2; Fig 3B; 155pp; English.
XX
CC The invention provides a cancer-associated epitope comprising two
CC separate polypeptides, a cytokeratin 8 polypeptide and a cytokeratin 18
CC polypeptide. Vaccine composition of the invention is useful for treating
CC or preventing colon adenocarcinoma, ovarian adenocarcinoma, renal
CC adenocarcinoma, mammary adenocarcinoma, lung adenocarcinoma, pancreatic
CC adenocarcinoma or non-seminoma testis carcinoma. The invention is also
CC useful for preparing a medicament for treating or preventing cancer in a
CC mammal. The present sequence is human cytokeratin K18 N-terminal peptide
XX
SQ Sequence 15 AA;
Query Match 3.6%; Score 76; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 STSFRGGMGSGGLAT 65
Db 1 STSFRGGMGSGGLAT 15

RESULT 11
ADJ81155
ID ADJ81155 standard; peptide; 27 AA.

XX AC ADJ81155;
XX DT 06-MAY-2004 (first entry)
XX DE Self-assembling polymer peptide #10.
XX KW miniblock polymer; optical response; mid infrared wavelength range;
KW long range ordered fluid; liquid crystals; nanolithography;
KW IR-sensitive device; IR sensor; IR filter; night telescope;
KW thermosensitive detector; non-linear chromophore; biomaterial;
KW tissue engineering scaffold; ferroelectric material; artificial muscle;
KW switching device.

XX OS Synthetic.

XX PN WO2003056297-A2.

XX PD 10-JUL-2003.

XX PF 02-OCT-2002; 2002WO-US031375.

XX PR 02-OCT-2001; 2001US-0326743P.

XX PR 04-JUN-2002; 2002US-0385809P.

XX PA (TUFT) TUFTS COLLEGE.

XX PI Valluzzi R, Kaplan DL;

XX DR WPI; 2003-671392/63.

XX PT Miniblock polymer useful in nanolithographic processes, has self-fabricating block and solubilizing block, has glycine and self-fabricates to form 3-dimensional materials having long-range orders, in solution.

XX PS Claim 29; SEQ ID NO 10; 91pp; English.

XX CC The invention relates to a miniblock polymer (I) comprising a self-fabricating block and a solubilizing block, a block for triggering self-fabrication by external or environmental conditions, and a block for incorporating turns in the polymer or for providing sites for chemical modifications, has a molecular weight of 1000-30000 and, in solution, can self-fabricate to form 3-dimensional material having long-range order, and where (I) has glycine content of at least 20%. (I) is useful for controlled delivery of a drug which involves incorporating a drug within (I), and administering the self-fabricating material incorporating the drug. The drug is incorporated within layers of the self-fabricating material. (II) is useful for modifying the optical response of a device in the near to mid infrared wavelength range which involves applying (II) to the surface of the device (all claimed). (I) is useful for preparing long range ordered fluids (i.e., liquid crystals) in a variety of phases or forms, which can then undergo very specific structural transitions to form rigid materials. (I) can be used as structural tissue implants, in liquid crystal displays, and for producing high-performance composites. (I) is useful for preparing chemically patterned templates with either general features are specific features. (I) is also useful in nanolithographic processes. A self-fabricated structure containing the polymer (II) is useful for modifying and improving performance of IR-sensitive devices, IR sensors, IR filters, night telescopes and thermosensitive detectors. (II) can be used for preparing films, membranes or coatings that absorb specific wavelengths of infrared radiation, and in optical applications e.g., as matrices to align non-linear chromophores which are useful in creating materials for second order non-linear optics. (II) is useful in hydrogen catalysis, as coatings for biomaterials, scaffolds for tissue engineering, ferroelectric materials, artificial muscles, switching devices, etc. This sequence represents a peptide used in the method of the invention.

XX SQ Sequence 27 AA;

Query Match 3.1%; Score 66; DB 7; Length 27;
Best Local Similarity 45.9%; Pred. No. 7.6e+02;
Matches 17; Conservative 3; Mismatches 5; Indels 12; Gaps 2;

QY 38 GAGSGSRISVSRSTSPRGGMGGL-ATGIAGGLAG 73
Db 2 GAGGGGT-----GGLGGGAGAGGLGGGGAG 27

RESULT 12

ADJ81362

ID ADJ81362 standard; peptide; 27 AA.

XX AC ADJ81362;

XX DT 06-MAY-2004 (first entry)

XX DE Self-assembling polymer peptide #217.

XX KW miniblock polymer; optical response; mid infrared wavelength range;
KW long range ordered fluid; liquid crystals; nanolithography;
KW IR-sensitive device; IR sensor; IR filter; night telescope;
KW thermosensitive detector; non-linear chromophore; biomaterial;
KW tissue engineering scaffold; ferroelectric material; artificial muscle;
KW switching device.

XX OS Synthetic.

XX PN WO2003056297-A2.

XX PD 10-JUL-2003.

XX PF 02-OCT-2002; 2002WO-US031375.

XX PR 02-OCT-2001; 2001US-0326743P.

XX PR 04-JUN-2002; 2002US-0385809P.

XX PA (TUFT) TUFTS COLLEGE.

XX PI Valluzzi R, Kaplan DL;

XX DR WPI; 2003-671392/63.

XX PT Miniblock polymer useful in nanolithographic processes, has self-fabricating block and solubilizing block, has glycine and self-fabricates to form 3-dimensional materials having long-range orders, in solution.

XX PS Disclosure; SEQ ID NO 218; 91pp; English.

XX CC The invention relates to a miniblock polymer (I) comprising a self-fabricating block and a solubilizing block, a block for triggering self-fabrication by external or environmental conditions, and a block for incorporating turns in the polymer or for providing sites for chemical modifications, has a molecular weight of 1000-30000 and, in solution, can self-fabricate to form 3-dimensional material having long-range order, and where (I) has glycine content of at least 20%. (I) is useful for controlled delivery of a drug which involves incorporating a drug within (I), and administering the self-fabricating material incorporating the drug. The drug is incorporated within layers of the self-fabricating material. (II) is useful for modifying the optical response of a device in the near to mid infrared wavelength range which involves applying (II) to the surface of the device (all claimed). (I) is useful for preparing long range ordered fluids (i.e., liquid crystals) in a variety of phases or forms, which can then undergo very specific structural transitions to form rigid materials. (I) can be used as structural tissue implants, in liquid crystal displays, and for producing high-performance composites. (I) is useful for preparing chemically patterned templates with either general features are specific features. (I) is also useful in nanolithographic processes. A self-fabricated structure containing the polymer (II) is useful for modifying and improving performance of IR-sensitive devices, IR sensors, IR filters, night telescopes and thermosensitive detectors. (II) can be used for preparing films, membranes or coatings that absorb specific wavelengths of infrared

CC radiation, and in optical applications e.g., as matrices to align non-
CC linear chromophores which are useful in creating materials for second
CC order non-linear orbits. (II) is useful in hydrogen catalysis, as
CC coatings for biomaterials, scaffolds for tissue engineering,
CC ferroelectric materials, artificial muscles, switching devices, etc. This
CC sequence represents a peptide used in the method of the invention.
XX
SQ Sequence 27 AA;

Query Match 3.1%; Score 66; DB 7; Length 27;
Best Local Similarity 45.9%; Pred. No. 7.6e+02;
Matches 17; Conservative 3; Mismatches 5; Indels 12; Gaps 2;

QY 38 GAGGSGSRISVSRSTSFRCGMSGGL-ATGIAGGLAG 73
||| |:
Db 2 GAGGGGT-----GGLSGGAGAGLGGGGAG 27

RESULT 13
ADJ81232
ID ADJ81232 standard; peptide; 27 AA.
XX
AC ADJ81232;
XX
DT 06-MAY-2004 (first entry)
XX
DE Self-assembling polymer peptide #87.

XX miniblock polymer; optical response; mid infrared wavelength range;
KW long range ordered fluid; liquid crystals; nanolithography;
KW IR-sensitive device; IR sensor; IR filter; night telescope;
KW thermosensitive detector; non-linear chromophore; biomaterial;
KW tissue engineering scaffold; ferroelectric material; artificial muscle;
KW switching device.

XX Synthetic.

XX WO2003056297-A2.

XX 10-JUL-2003.

XX 02-OCT-2002; 2002WO-US031375.

XX 02-OCT-2001; 2001US-0326743P.

XX 04-JUN-2002; 2002US-0385809P.

XX (TUFT) TUFTS COLLEGE.

XX Valluzzi R, Kaplan DL;

XX WPI; 2003-671392/63.

XX Miniblock polymer useful in nanolithographic processes, has self-
PT fabricating block and solubilizing block, has glycine and self-fabricates
PT to form 3-dimensional materials having long-range orders, in solution.

PS Claim 32; SEQ ID NO 87; 91pp; English.

XX The invention relates to a miniblock polymer (I) comprising a self-
CC fabricating block and a solubilizing block, a block for triggering self-
CC fabrication by external or environmental conditions, and a block for
CC incorporation turns in the polymer or for providing sites for chemical
CC modifications, has a molecular weight of 1000-300000 and, in solution,
CC can self-fabricate to form 3-dimensional material having long-range
CC order, and where (I) has glycine content of at least 20%. (I) is useful
CC for controlled delivery of a drug which involves incorporating a drug
CC within (I), and administering the self-fabricating material incorporating
CC the drug. The drug is incorporated within layers of the self-fabricating
CC material. (II) is useful for modifying the optical response of a device
CC in the near to mid infrared wavelength range which involves applying (II)
CC to the surface of the device (all claimed). (I) is useful for preparing
CC long range ordered fluids (i.e., liquid crystals) in a variety of phases
CC or forms, which can then undergo very specific structural transitions to

CC form rigid materials. (I) can be used as structural tissue implants, in
CC liquid crystal displays, and for producing high-performance composites.
CC (I) is useful for preparing chemically patterned templates with either
CC general features or specific features. (I) is also useful in
CC nanolithographic processes. A self-fabricated structure containing the
CC polymer (II) is useful for modifying and improving performance of IR-
CC sensitive devices, IR sensors, IR filters, night telescopes and
CC thermosensitive detectors. (II) can be used for preparing films,
CC membranes or coatings that absorb specific wavelengths of infrared
CC radiation, and in optical applications e.g., as matrices to align non-
CC linear chromophores which are useful in creating materials for second
CC order non-linear orbits. (II) is useful in hydrogen catalysis, as
CC coatings for biomaterials, scaffolds for tissue engineering,
CC ferroelectric materials, artificial muscles, switching devices, etc. This
CC sequence represents a peptide used in the method of the invention.
XX

SQ Sequence 27 AA;

Query Match 3.1%; Score 66; DB 7; Length 27;
Best Local Similarity 45.9%; Pred. No. 7.6e+02;
Matches 17; Conservative 3; Mismatches 5; Indels 12; Gaps 2;

QY 38 GAGGSGSRISVSRSTSFRCGMSGGL-ATGIAGGLAG 73
||| |:
Db 2 GAGGGGT-----GGLSGGAGAGLGGGGAG 27

RESULT 14

ADJ81293

ID ADJ81293 standard; peptide; 27 AA.

XX ADJ81293;

XX 06-MAY-2004 (first entry)

XX Self-assembling polymer peptide #148.

XX miniblock polymer; optical response; mid infrared wavelength range;
KW long range ordered fluid; liquid crystals; nanolithography;
KW IR-sensitive device; IR sensor; IR filter; night telescope;
KW thermosensitive detector; non-linear chromophore; biomaterial;
KW tissue engineering scaffold; ferroelectric material; artificial muscle;
KW switching device.

XX Synthetic.

XX WO2003056297-A2.

XX 10-JUL-2003.

XX 02-OCT-2002; 2002WO-US031375.

XX 02-OCT-2001; 2001US-0326743P.

XX 04-JUN-2002; 2002US-0385809P.

XX (TUFT) TUFTS COLLEGE.

XX Valluzzi R, Kaplan DL;

XX WPI; 2003-671392/63.

XX Miniblock polymer useful in nanolithographic processes, has self-
PT fabricating block and solubilizing block, has glycine and self-fabricates
PT to form 3-dimensional materials having long-range orders, in solution.

PS Disclosure; SEQ ID NO 149; 91pp; English.

XX The invention relates to a miniblock polymer (I) comprising a self-
CC fabricating block and a solubilizing block, a block for triggering self-
CC fabrication by external or environmental conditions, and a block for
CC incorporating turns in the polymer or for providing sites for chemical
CC modifications, has a molecular weight of 1000-300000 and, in solution,
CC can self-fabricate to form 3-dimensional material having long-range

CC order, and where (I) has glycine content of at least 20%. (I) is useful
CC for controlled delivery of a drug which involves incorporating a drug
CC within (I), and administering the self-fabricating material incorporating
CC the drug. The drug is incorporated within layers of the self-fabricating
CC material. (II) is useful for modifying the optical response of a device
CC in the near to mid infrared wavelength range which involves applying (II)
CC to the surface of the device (all claimed). (I) is useful for preparing
CC long range ordered fluids (i.e., liquid crystals) in a variety of phases
CC or forms, which can then undergo very specific structural transitions to
CC form rigid materials. (I) can be used as structural tissue implants, in
CC liquid crystal displays, and for producing high-performance composites.
CC (I) is useful for preparing chemically patterned templates with either
CC general features are specific features. (I) is also useful in
CC nanolithographic processes. A self-fabricated structure containing the
CC polymer (II) is useful for modifying and improving performance of IR-
CC sensitive devices, IR sensors, IR filters, night telescopes and
CC thermosensitive detectors. (II) can be used for preparing films,
CC membranes or coatings that absorb specific wavelengths of infrared
CC radiation, and in optical applications e.g., as matrices to align non-
CC linear chromophores which are useful in creating materials for second
CC order non-linear optics. (II) is useful in hydrogen catalysis, as
CC coatings for biomaterials, scaffolds for tissue engineering,
CC ferroelectric materials, artificial muscles, switching devices, etc. This
CC sequence represents a peptide used in the method of the invention.
XX
SQ Sequence 27 AA;

Query Match 3.1%; Score 66; DB 7; Length 27;
Best Local Similarity 45.9%; Pred. NO. 7.6e+02;
Matches 17; Conservative 3; Mismatches 5; Indels 12; Gaps 2;
QY 38 GAGSGSRISVSRSTFRGGMGGGL-ATGIAGGLAG 73
Db 2 GAGGGGT-----GGLGSGGAGAGLGGGGAG 27

RESULT 15
AAU68081
ID AAU68081 standard; peptide; 13 AA.
XX
AC AAU68081;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human Breast cancer-associated protein isoform, BPI-309 peptide #6.
XX
KW Human; Breast cancer-associated protein isoform; breast cancer;
KW immunogen; cytostatic; BPI; tryptic digest peptide.
XX
OS Homo sapiens.
XX
PN WO200171357-A2.
XX
PD 27-SEP-2001.
XX
PF 20-MAR-2001; 2001WO-GB001219.
XX
PR 20-MAR-2000; 2000GB-00006695.
PR 24-MAR-2000; 2000GB-00007265.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMCAC, O'hare MJ, Page MJ, Parekh RB, Waterfield MD;
XX
XX WPI; 2001-611532/70.
XX
PT Identifying proteins for clinical screening, diagnosis and prognosis of
PT breast cancer, comprises detecting Breast Cancer-Associated Protein
PT Isoforms (BPIs) using two-dimensional electrophoresis.
XX
PS Claim 9; Page 44; 197pp; English.
XX
CC The invention relates to diagnosing, determining the stage or severity,

CC or identifying the risk of a subject developing cancer (especially breast
CC cancer), or monitoring the effect of therapy on a subject with cancer,
CC comprising analysing a test sample using two-dimensional electrophoresis
CC and detecting Breast Cancer-Associated Protein Isoforms (BPIs). The
CC methods disclosed are used for the diagnosis and prognosis of breast
CC cancer, for determining the severity of breast cancer, and for
CC identifying a subject at risk of developing breast cancer, and monitoring
CC the effect of therapy administered to a subject. Antibodies raised
CC against the binding domain of a BPI, or a nucleic acid that inhibits the function
CC of a BPI can be incorporated into a pharmaceutical composition for
CC treating or preventing breast cancer. The methods use sensitive and
CC specific biomarkers provide early diagnosis of breast cancer, and the
CC compositions are more potent, specific, and has a more rapid effect with
CC fewer side effects than other prior art methods. The present sequence is
CC a tryptic digest peptide from a BPI of the invention
XX
SQ Sequence 13 AA;

Query Match 3.0%; Score 64; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. NO. 4.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 SLGSVQAPSYGAR 27
Db 1 SLGSVQAPSYGAR 13

Search completed: December 14, 2004, 09:29:37
Job time : 153 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31 ; Search time 7.11354 Seconds
(without alignments)
727.960 Million cell updates/sec

Title: US-10-026-001-3
Perfect score: 51
Sequence: 1 FLDKGMHYV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	74.5	141	Q7RJX9	Q7rx9 plasmodium
2	38	74.5	182	Q6D008	Q6d008 erwinia car
3	38	74.5	208	3MCH_RHILO	Q98dr6 rhizobium l
4	38	74.5	280	Q8II87	Q8ii87 plasmodium
5	38	74.5	1209	Q6Y5E2	Q6y5e2 oryctolagus
6	38	74.5	1209	AAO47084	Aao47084 oryctolag
7	37	72.5	162	Q98TN1	Q98tn1 platichthys
8	37	72.5	215	1 VH22 MYXVL	Q9q8n5 myxoma viru
9	37	72.5	216	CTRD_NEIMA	P57013 neisseria m
10	37	72.5	216	CTRD_NEIMB	P32016 neisseria m
11	37	72.5	222	Q88DL3	Q88dl3 pseudomonas
12	37	72.5	239	1 3MCH_PSEAE	Q9hxl7 pseudomonas
13	37	72.5	285	Q74I29	Q74iz9 lactobacill
14	37	72.5	285	AAO9132	Aao9132 lactobaci
15	37	72.5	326	Q7PXJ7	Q7pxj7 anopheles g
16	37	72.5	401	Q8ILK1	Q8ilk1 plasmodium
17	37	72.5	470	1 SYC_SULSO	Q97we6 sulfolobus
18	37	72.5	471	1 SYC_SULTO	Q96yc3 sulfolobus
19	37	72.5	652	Q68071	Q68071 rhodobacter
20	37	72.5	1065	1 CERU_HUMAN	P00450 homo sapien
21	37	72.5	4799	2 Q7PQW0	Q7pqw0 anopheles g
22	36	70.6	139	Q914M7	Q914m7 sulfolobus
23	36	70.6	208	1 209E_MOUSE	Q91zw7 mus musculu
24	36	70.6	272	Q8GEM1	Q8gem1 pseudomonas
25	36	70.6	272	Q88PL2	Q88pl2 pseudomonas
26	36	70.6	397	2 Q99PR7	Q99pr7 cavia porce
27	36	70.6	430	2 Q8BIY3	Q8biy3 mus musculu
28	36	70.6	480	2 Q88QY7	Q88qy7 pseudomonas
29	36	70.6	502	2 Q8ALM8	Q8alm8 bacteroides
30	36	70.6	570	2 Q8Y80	Q8yy80 anabaena sp
31	36	70.6	673	2 Q6CE11	Q6cel1 yarrowia li

32	36	70.6	907	2	Q74F46	Q74f46 geobacter s
33	36	70.6	907	2	AAR34093	Aar34093 geobacter
34	36	70.6	967	1	SYA_BOMMO	P21894 bombyx mori
35	36	70.6	1032	2	Q8AXZ4	Q8axz4 brachydanio
36	36	70.6	1199	2	Q6E5V9	Q6e5v9 gecarcinus
37	35	68.6	130	2	Q73J87	Q73j87 treponema d
38	35	68.6	130	2	AAS13206	Aas13206 treponema
39	35	68.6	172	2	Q8KJQ3	Q8kjq3 escherichia
40	35	68.6	192	2	Q6MGX6	Q6mgx6 bdellovibri
41	35	68.6	192	2	CAE81153	CaE81153 bdellovib
42	35	68.6	233	2	Q7XKP9	Q7xkp9 cryza sativ
43	35	68.6	237	2	Q93H72	Q93h72 streptomyce
44	35	68.6	254	2	Q73L74	Q73l74 treponema d
45	35	68.6	254	2	Q73LH2	Q73lh2 treponema d

ALIGNMENTS

RESULT 1
Q7RJX9 PRELIMINARY; PRT; 141 AA.
AC Q7RJX9; PRELIMINARY; PRT; 141 AA.
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY03126;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AABL0100082; EAA22658.1; --
DR EMBL; AABL0100082; EAA22658.1; --
KW Hypothetical protein.
SQ SEQUENCE 141 AA; 16992 MW; 3F38737DDDD35B830 CRC64;

Query Match 74.5%; Score 38; DB 2; Length 141;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLDKGMHY 8
||:||||
Db 87 FLNEGMY 94

RESULT 2
Q6D008 PRELIMINARY; PRT; 182 AA.
AC Q6D008;
DT 01-OCT-2004 (Tremblrel. 28, Created)
DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Putative transferase.
GN ORFNames=ECA3993;

OS Erwinia carotovora subsp. atroseptica SCRI1043.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OX NCBI_TaxID=218491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SCRI1043;
 RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,
 RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
 RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
 RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
 RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
 RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.,
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX950851; CAG76890.1; -
 KW Transference.
 SQ SEQUENCE 182 AA; 19783 MW; 46E19BFE565BF67 CRC64;
 Query Match 74.5%; Score 38; DB 2; Length 182;
 Best Local Similarity 75.0%; Pred. No. 24;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LDKGHMYV 9
 DB 137 LEKGHLV 144
 RESULT 3
 3MCH_RHILO STANDARD; PRT; 208 AA.
 AC Q98DR6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Putative 3-methyladenine DNA glycosylase (EC 3.2.2.-).
 GN OrderedLocusNames=ml14583;
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFP303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 CC -1- SIMILARITY: Belongs to the DNA glycosylase MPG family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP003004; BAB51204.1; -
 DR HSSP; P29372; 1F60.
 DR HAMAP; MF_00527; -; 1.
 DR InterPro; IPR011034; FMT_C like.
 DR InterPro; IPR003180; PurDNA_glycylase.
 DR Pfam; PF02245; PurDNA_glyco; 1.
 DR ProDom; PD009649; PurDNA_glycylase; 1.
 DR TIGRFAMs; TIGR00567; 3mg; 1.
 KW Complete proteome; DNA repair; Hydrolase; Hypothetical protein.
 SQ SEQUENCE 208 AA; 22771 MW; 9972F578411E85E CRC64;

Query Match 74.5%; Score 38; DB 1; Length 208;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FLDKGMVYV 9
 DB 77 FLERGHAYV 85
 RESULT 4
 Q8II87 PRELIMINARY; PRT; 280 AA.
 ID Q8II87;
 AC Q8II87;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=PF11_0287;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2255705; PubMed=12368864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.S., Nene V., Shallow S.J., Su B., Peterson J., Angiuoli S.,
 RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 RT falciparum."
 RL Nature 419:498-511(2002).
 DR EMBL; AB014840; AAN35871.1; -
 DR InterPro; IPR001251; CRAL_TRIO C.
 DR InterPro; IPR011074; Sec14p_like N.
 DR Pfam; PF00650; CRAL_TRIO; 1.
 DR SMART; SM00516; SEC14; 1.
 DR PROSITE; PS0191; CRAL_TRIO; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 280 AA; 33118 MW; BDC8D1DC0F557A55 CRC64;
 Query Match 74.5%; Score 38; DB 2; Length 280;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FLDKGMVYV 9
 DB 117 FLDKGYCYI 125
 RESULT 5
 Q6Y5E2 PRELIMINARY; PRT; 1209 AA.
 ID Q6Y5E2;
 AC Q6Y5E2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Endothelial nitric oxide synthase NOS3.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Vascular endothelium;
 RA Rachlis A.C., Wang Y., Miller T.L., Robb G.B., Cybulsky M.I.,
 RA Marsden P.A.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY179960; AAO47084.1; -

```
DR InterPro; IPR003097; FAD_binding.  
DR InterPro; IPR001094; Flavodoxin_like.  
DR InterPro; IPR008254; Flav_nitox_synth.  
DR InterPro; IPR001709; FPN_Cyt_redctse.  
DR InterPro; IPR004030; NO_synthase.  
DR InterPro; IPR001433; Oxred_FAD/NAD(P).  
DR Pfam; PF00667; FAD_binding_1; 1.  
DR Pfam; PF00258; Flavodoxin_1; 1.  
DR Pfam; PF00175; NAD_binding_1; 1.  
DR Pfam; PF02898; NO_synthase; 1.  
DR PRINTS; PR00371; FPNCR.  
DR PROSITE; PS00902; FLAVODOXIN LIKE; 1.  
DR PROSITE; PS00001; NOS; UNKNOWN_1.  
SQ SEQUENCE 1209 AA; 133839 MW; 23EF203E2A7EF14A CRC64;  
  
Query Match 74.5%; Score 38; DB 2; Length 1209;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 LDKGHMYV 9  
Db 1112 LERGHMYV 1119  
|:|||||  
|:|||||  
  
RESULT 6  
AAO47084 PRELIMINARY; PRT; 1209 AA.  
ID AAO47084  
AC AAO47084;  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE Endothelial nitric oxide synthase NOS3.  
OS Oryctolagus cuniculus (Rabbit)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Vascular endothelium;  
RA Rachlis A.C., Wang Y., Miller T.L., Robb G.B., Cybulsky M.I.,  
RA Marsden P.A.;  
RT "Cloning and Characterization of Rabbit Endothelial Nitric Oxide  
RT Synthase."  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY179960; AAO47084.1; -.  
SQ SEQUENCE 1209 AA; 133839 MW; 23EF203E2A7EF14A CRC64;  
  
Query Match 74.5%; Score 38; DB 2; Length 1209;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 LDKGHMYV 9  
Db 1112 LERGHMYV 1119  
|:|||||  
|:|||||  
  
RESULT 7  
Q98TN1 PRELIMINARY; PRT; 162 AA.  
ID Q98TN1  
AC Q98TN1;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Paraoxonase (Arylesterase) (Fragment).  
GN Name=pon;  
OS Platicthys flesus (European flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidae; Pleuronectidae; Platicthys.  
OX NCBI_TaxID=8260;  
RN [1]
```

```
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=22827201; PubMed=12946615;  
RA Williams T.D., Gensberg K., Minchin S.D., Chipman J.K.;  
RT "A DNA expression array to detect toxic stress response in European  
RT flounder (Platichthys flesus).";  
RL Aquat. Toxicol. 65:141-157(2003).  
DR EMBL; AJ292086; FAD28869.1; -.  
DR GO; GO:0004064; F:arylesterase activity; IEA.  
DR InterPro; IPR002640; Arylesterase.  
DR InterPro; IPR011045; N2O reductase_N.  
DR Pfam; PF01731; Arylesterase; 1.  
DR PRINTS; PR01785; PARAOXONASE.  
FT NON_TER 1  
FT NON_TER 162  
SQ SEQUENCE 162 AA; 17917 MW; D3FA421BAFF6354C CRC64;  
  
Query Match 72.5%; Score 37; DB 2; Length 162;  
Best Local Similarity 85.7%; Pred. No. 34;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 DKGHMYV 9  
Db 52 DKGHYV 58  
|:|||||  
|:|||||  
  
RESULT 8  
VH22_MYXVL STANDARD; PRT; 215 AA.  
ID VH22_MYXVL  
AC Q9Q8N5;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Probable host range protein 2-2.  
GN Name=M063R;  
OS Myxoma virus (strain Lausanne) (MYXV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Leporipoxvirus.  
OX NCBI_TaxID=31530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20032073; PubMed=10562494;  
RA Cameron C., Hota-Mitchell S., Chen L., Barrett J., Cao J.X.,  
RA Macaulay C., Willer D., Evans D., McFadden G.;  
RT "The complete DNA sequence of myxoma virus."  
RL Virology 264:298-318(1999).  
CC -!- FUNCTION: Plays a role for multiplication of the virus in  
CC different cell types (By similarity).  
CC -!- SIMILARITY: Belongs to the poxviruses C7 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; AF170726; AAF14951.1; -.  
DR InterPro; IPR004967; Pox_C7_P8A.  
DR Pfam; PF03287; Pox_C7_P8A; 1.  
DR PIRSF; PIRSF003779; VAC_C7L; 1.  
SQ SEQUENCE 215 AA; 24578 MW; 9D5D3CF8FCB74B27 CRC64;  
  
Query Match 72.5%; Score 37; DB 1; Length 215;  
Best Local Similarity 75.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 FLDKGHMY 8  
Db 173 FMDKGY 180  
|:|||||  
|:|||||
```

```

RESULT 9
CTRD_NEIMA
ID_CTRD_NEIMA STANDARD; PRT; 216 AA.
AC PS7013;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Capsule polysaccharide export ATP-binding protein ctd (EC 3.6.3.38)
DE (Capsular-polysaccharide-exporting ATPase).
GN Name=ctrD; OrderedLocusNames=NMA0195;
OS Neisseria meningitidis (serogroup A);
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Murgall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
CC -!- FUNCTION: Putative ATP-binding protein, and an energy-coupling
CC component of capsule polysaccharide export apparatus.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + capsular polysaccharide(In) =
CC ADP + phosphate + capsular polysaccharide(Out).
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL162752; CAB83509.1; -.
CC PIR; F82013; F82013.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC Pfam; PF00005; ABC_tran; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Bacterial capsule; Complete proteome; Hydrolase;
KW Inner membrane; Polysaccharide transport; Transport.
FT NP BIND 38 45 ATP (By similarity).
SQ SEQUENCE 216 AA; 24844 MW; F265630711D674BF CRC64;

Query Match 72.5%; Score 37; DB 1; Length 216;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
Db 194 LEKGHMY 200

RESULT 10
CTRD_NEIMA
ID_CTRD_NEIMA STANDARD; PRT; 216 AA.
AC P32016;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Capsule polysaccharide export ATP-binding protein ctd (EC 3.6.3.38)
DE (Capsular-polysaccharide-exporting ATPase).
GN Name=ctrD; OrderedLocusNames=NMA0195;
OS Neisseria meningitidis (serogroup A);
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Murgall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
CC -!- FUNCTION: Putative ATP-binding protein, and an energy-coupling
CC component of capsule polysaccharide export apparatus.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + capsular polysaccharide(In) =
CC ADP + phosphate + capsular polysaccharide(Out).
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL162752; CAB83509.1; -.
CC PIR; F82013; F82013.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC Pfam; PF00005; ABC_tran; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Bacterial capsule; Complete proteome; Hydrolase;
KW Inner membrane; Polysaccharide transport; Transport.
FT NP BIND 38 45 ATP (By similarity).
SQ SEQUENCE 216 AA; 24844 MW; F265630711D674BF CRC64;

Query Match 72.5%; Score 37; DB 1; Length 216;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
Db 194 LEKGHMY 200

```

```

GN Name=ctrD; OrderedLocusNames=NMB00074;
OS Neisseria meningitidis (serogroup B);
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B1940 / Serogroup B;
RX MEDLINE=92065822; PubMed=1659649;
RA Frosch M., Edwards U., Bousset K., Krause B., Weisgerber C.;
RT "Evidence for a common molecular origin of the capsule gene loci in
RT Gram-negative bacteria expressing group II capsular polysaccharides."
RL Mol. Microbiol. 5:1251-1263(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Citterone H., Clark E.B., Cotton M.D., Uitterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizzi M.,
RA Grandi G., Sun L., Smith K.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
CC -!- FUNCTION: Putative ATP-binding protein, and an energy-coupling
CC component of capsule polysaccharide export apparatus.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + capsular polysaccharide(In) =
CC ADP + phosphate + capsular polysaccharide(Out).
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M57677; AAA25453.1; -.
CC EMBL; AE002366; AAF40541.1; -.
CC PIR; H81241; H81241.
CC PIR; S15223; S15223.
CC TIGR; NMB0074; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC Pfam; PF00005; ABC_tran; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Bacterial capsule; Complete proteome; Hydrolase;
KW Inner membrane; Polysaccharide transport; Transport.
FT NP BIND 38 45 ATP (By similarity).
FT CONFLICT 10 17 RYLTRGGM -> QYQMRGGM (in Ref. 1).
FT CONFLICT 22 22 H -> D (in Ref. 1).
FT CONFLICT 25 29 SFKWE -> NFSLQ (in Ref. 1).
FT CONFLICT 34 34 I -> V (in Ref. 1).
FT CONFLICT 48 48 I -> V (in Ref. 1).
FT CONFLICT 59 59 T -> S (in Ref. 1).
SQ SEQUENCE 216 AA; 24801 MW; 5608DEB1CABBC909 CRC64;

Query Match 72.5%; Score 37; DB 1; Length 216;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
Db 194 LEKGHMY 200

```

```

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- SIMILARITY: Belongs to the DNA glycosylase MPG family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE004818; AAC07397.1; -.
DR PIR; G83143; G83143.
DR HSSP; P29372; 1F60.
DR HAVAP; MF 00527; -.
DR InterPro; IPR011034; FMT C like.
DR InterPro; IPR003180; PurDNA_glycosylase.
DR Pfam; PF02245; Pur DNA glyco; 1.
DR ProDom; PD009649; PurDNA_glycosylase; 1.
KW Complete proteome; DNA repair; Hydrolase; Hypothetical protein.
SQ SEQUENCE 239 AA; 26949 MW; 802F0365A74D95D8 CRC64;

Query Match 72.5%; Score 37; DB 1; Length 239;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
Db 78 FLDGGHIYM 86

RESULT 13
Q741Z9 PRELIMINARY; PRT; 285 AA.
AC Q741Z9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LJ1312;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwaalen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017204; AAS09132.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 285 AA; 33529 MW; 88D2907845FE8423 CRC64;

Query Match 72.5%; Score 37; DB 2; Length 285;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 7
Db 106 FLDKGHMYV 112

RESULT 14
AAS09132

```

ID AAS09132 PRELIMINARY; PRT; 285 AA.
AC AAS09132;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN LJ1312.
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwaalen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.,
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017204; AAS09132.1; -.
KW Hypothetical protein.
SQ SEQUENCE 285 AA; 33529 MW; 88D2907845FE8423 CRC64;

Query Match 72.5%; Score 37; DB 2; Length 285;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLDKGHM 7
DB 106 FLDKGHL 112

RESULT 15
Q7PXJ7 PRELIMINARY; PRT; 326 AA.
AC Q7PXJ7;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE AGC12206 (Fragment).
GN Name=agCG48004; ORFNames=ENSANGG00000013252;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008987; EAA01751.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000868; Iscrsm_hydrolase.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF00857; Isochorismatase; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36936 MW; 18C5E9F85F47A196 CRC64;

Query Match 72.5%; Score 37; DB 2; Length 326;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 DKGHMYV 9

Db 20 DKGHIYV 26
Search completed: December 14, 2004, 09:19:06
Job time : 10.1135 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31 ; Search time 6.93668 Seconds
(without alignments)
465.433 Million cell updates/sec

Title: US-10-026-001-5

Perfect score: 47

Sequence: 1 FLDKAHMEV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	47	100.0	9	5 ABB81089	ABB81089 Human cyt
2	36	76.6	9	5 ABB81088	ABB81088 Human cyt
3	34	72.3	377	4 ABO3039	ABO3039 Novel hum
4	34	72.3	422	7 ABO5926	ABO5926 Klebsiell
5	34	72.3	429	4 AAB92560	AAB92560 Human pro
6	34	72.3	429	4 AAG81322	AAG81322 Human AFP
7	34	72.3	429	5 AAU81499	AAU81499 Mouse int
8	34	72.3	429	5 AAU81512	AAU81512 Human int
9	34	72.3	429	8 ADO48145	ADO48145 Human HIP
10	34	72.3	771	4 ABG24971	ABG24971 Novel hum
11	34	72.3	1683	4 ABG24799	ABG24799 Novel hum
12	33	70.2	151	5 ABR01783	ABR01783 Human bre
13	33	70.2	220	2 AAY59860	AAY59860 Human nor
14	33	70.2	364	2 AAW62047	AAW62047 Streptoco
15	33	70.2	366	6 ABB98965	ABB98965 Human KIA
16	33	70.2	740	2 AAW62045	AAW62045 Streptoco
17	33	70.2	740	4 AAU37890	AAU37890 Streptoco
18	33	70.2	740	4 AAM01059	AAM01059 CFE 62 pr
19	33	70.2	740	6 ABU02112	ABU02112 S. pneumo
20	33	70.2	740	6 ABU46184	ABU46184 Protein e
21	33	70.2	740	8 ADK46710	ADK46710 Streptoco
22	33	70.2	740	8 ADM92202	ADM92202 S pneumon
23	32	68.1	143	7 ABO77628	ABO77628 Pseudomon
24	32	68.1	190	6 ABU26391	ABU26391 Protein e
25	32	68.1	198	6 ABU41146	ABU41146 Protein e

26	32	68.1	212	7 ADF04417	ADF04417 Bacterial
27	32	68.1	326	6 ABU23850	ABU23850 Protein e
28	32	68.1	336	7 ABO82607	ABO82607 Pseudomon
29	32	68.1	440	3 AAB18146	AAB18146 Plasmodiu
30	32	68.1	455	7 ADM25401	ADM25401 Hyperther
31	32	68.1	467	5 ABP55691	ABP55691 Human HCC
32	32	68.1	494	2 AAY34797	AAY34797 Chlamydia
33	32	68.1	507	6 ABU21060	ABU21060 Protein e
34	32	68.1	527	5 ABB90589	ABB90589 Chlamydia
35	32	68.1	527	5 ABP61995	ABP61995 C. pneumo
36	32	68.1	582	6 ABU20039	ABU20039 Protein e
37	32	68.1	589	6 ABU22663	ABU22663 Protein e
38	32	68.1	594	6 ABP78711	ABP78711 N. gonorr
39	32	68.1	594	6 ABU37090	ABU37090 Protein e
40	32	68.1	594	6 ABU38019	ABU38019 Protein e
41	32	68.1	594	8 ADR08296	ADR08296 Neisseria
42	32	68.1	599	2 AAR22008	AAR22008 P64KD pro
43	32	68.1	635	2 AAW75410	AAW75410 Fusion pr
44	32	68.1	635	2 AAW75409	AAW75409 Fusion pr
45	32	68.1	635	2 AAW75412	AAW75412 Fusion pr

ALIGNMENTS

RESULT 1
ABB81089
ID ABB81089 standard; peptide; 9 AA.
XX
AC ABB81089;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human cytokeratin 18 (CK18) compound 2 peptide.
XX
KW Antigen; cytokeratin 18; CK-18; immune response; cytostatic; vaccine;
KW gene therapy; cancer; human.
XX
OS Homo sapiens.
XX
PN WO200255555-A2.
XX
PD 18-JUL-2002.

XX
PE 21-DEC-2001; 2001WO-US049964.
XX
PR 21-DEC-2000; 2000US-0257820P.
XX
(GENZ) GENZYME CORP.
XX
PI Nicolette CA;
XX
DR WPI; 2002-619103/66.
XX
N-PSDB; ABN86612.
XX
Novel antigenic cytokeratin 18 compounds and peptides useful for inducing an immune response in a subject and for diagnosing a neoplastic condition or susceptibility to the condition of an animal cell or tissue.
XX
Claim 2; Page 63; 73pp; English.
XX
The invention relates to novel antigenic cytokeratin 18 (CK-18) compounds (I) and peptides useful for inducing an immune response in a subject. (I) is useful for inducing an immune response in a subject, by delivering (I) as a polynucleotide, in the context of an MHC molecule that presents the compound on the surface of an APC. (I) is useful for aiding in the diagnosis of the neoplastic condition or susceptibility to the condition of an animal cell or tissue; for generating antibodies which are useful for identifying and purifying polypeptides and APCs expressing the polypeptides. (I) serves as markers for the neoplastic phenotype. (I) that is covalently or non-covalently linked to molecules are useful in diagnostic methods, and for detecting or purifying antibodies. It is also useful as components of anti-cancer vaccines and to expand immune

CC effector cells that are specific for cells having differential, i.e.
CC aberrant, expression of antigenic CK-18. CK-18 proteins or antibodies are
CC useful for detecting, diagnosing or prognosing and monitoring the
CC progression, course or stage of CK-18 related cancers, or malignancies.
CC Host cells comprising one or more immunogenic ligands are useful for
CC inducing an immune response in a subject, and to expand a population of
CC immune effector cells such as tumour infiltrating lymphocytes which in
CC turn are useful in adoptive immunotherapies. Agents that modulate the
CC binding of CK-18 protein to its ligand are useful for treating disease,
CC especially cancer. The present sequence represents the human CK18
CC compound 2 peptide
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 47; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|||||
Db 1 FLDKAHMEV 9

RESULT 2
ABB81088
ID ABB81088 standard; peptide; 9 AA.

AC ABB81088;

DT 05-NOV-2002 (first entry)

XX Human cytokeratin 18 (CK18) compound 1 peptide.

DE Antigen; cytokeratin 18; CK-18; immune response; cytostatic; vaccine;
KW gene therapy; cancer; human.

OS Homo sapiens.

PN WO200255555-A2.

PD 18-JUL-2002.

PF 21-DEC-2001; 2001WO-US049964.

PR 21-DEC-2000; 2000US-0257820P.

PA (GENZ) GENZYME CORP.

PI Nicolette CA;

XX WPI; 2002-619103/66.

DR N-PSDB; ABN86611.

XX Novel antigenic cytokeratin 18 compounds and peptides useful for inducing
PT an immune response in a subject and for diagnosing a neoplastic condition
PT or susceptibility to the condition of an animal cell or tissue.

PS Claim 1; Page 63; 73pp; English.

XX The invention relates to novel antigenic cytokeratin 18 (CK-18) compounds
CC (I) and peptides useful for inducing an immune response in a subject. (I)
CC is useful for inducing an immune response in a subject, by delivering (I)
CC as a polynucleotide, in the context of an MHC molecule that presents the
CC compound on the surface of an APC. (I) is useful for aiding in the
CC diagnosis of the neoplastic condition or susceptibility to the condition
CC of an animal cell or tissue; for generating antibodies which are useful
CC for identifying and purifying polypeptides and APCs expressing the
CC polypeptides. (I) serves as markers for the neoplastic phenotype. (I)
CC that is covalently or non-covalently linked to molecules are useful in
CC diagnostic methods, and for detecting or purifying antibodies. It is also
CC useful as components of anti-cancer vaccines and to expand immune
CC effector cells that are specific for cells having differential, i.e.
CC aberrant, expression of antigenic CK-18. CK-18 proteins or antibodies are

CC useful for detecting, diagnosing or prognosing and monitoring the
CC progression, course or stage of CK-18 related cancers, or malignancies.
CC Host cells comprising one or more immunogenic ligands are useful for
CC inducing an immune response in a subject, and to expand a population of
CC immune effector cells such as tumour infiltrating lymphocytes which in
CC turn are useful in adoptive immunotherapies. Agents that modulate the
CC binding of CK-18 protein to its ligand are useful for treating disease,
CC especially cancer. The present sequence represents the human CK18
CC compound 1 peptide
XX
SQ Sequence 9 AA;

Query Match 76.6%; Score 36; DB 5; Length 9;
Best Local Similarity 77.8%; Pred. No. 1.7e+06;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|||||
Db 1 FLDKCHMYV 9

RESULT 3
ABG03039
ID ABG03039 standard; protein; 377 AA.

AC ABG03039;

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #3030.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS67226.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 33398; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 377 AA;

Query Match 72.3%; Score 34; DB 4; Length 377;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
 |||||
 Db 18 FLDKLNHNEI 26

RESULT 4

ABO65926
 ID ABO65926 standard; protein; 422 AA.

XX
 AC ABO65926;

XX
 DT 29-JUL-2004 (first entry)

XX
 DE Klebsiella pneumoniae polypeptide seqid 12443.

XX
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX
 OS Klebsiella pneumoniae.

XX
 PN US6610836-B1.

XX
 PD 26-AUG-2003.

XX
 PF 27-JAN-2000; 2000US-00489039.

XX
 PR 29-JAN-1999; 99US-0117747P.

XX
 PA (GENO-) GENOME THERAPEUTICS CORP.

XX
 PI Breton GL, Osborne M;

XX
 DR WPI; 2003-895346/82.
 DR N-PSDB; ACH99477.

XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.

XX
 PS Disclosure; SEQ ID NO 12443; 932pp; English.

XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention

XX
 SQ Sequence 422 AA;

Query Match 72.3%; Score 34; DB 7; Length 422;
 Best Local Similarity 62.5%; Pred. No. 1.7e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
 |||||
 Db 87 FIDKGHLE 94

RESULT 5

AAB92560
 ID AAB92560 standard; protein; 429 AA.

XX
 AC AAB92560;

XX
 DT 26-JUN-2001 (first entry)

XX
 DE Human protein sequence SEQ ID NO:10755.

XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX
 OS Homo sapiens.

XX
 PN EP1074617-A2.

XX
 PD 07-FEB-2001.

XX
 PF 28-JUL-2000; 2000EP-00116126.

XX
 PR 29-JUL-1999; 99JP-00248036.

XX
 PR 27-AUG-1999; 99JP-00300253.

XX
 PR 11-JAN-2000; 2000JP-00118776.

XX
 PR 02-MAY-2000; 2000JP-00183767.

XX
 PR 09-JUN-2000; 2000JP-00241899.

XX
 PA (HELI-) HELIX RES INST.

XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX
 DR WPI; 2001-318749/34.

XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.

XX
 PS Claim 8; SEQ ID NO 10755; 2537pp + Sequence Listing; English.

XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX
 SQ Sequence 429 AA;

Query Match 72.3%; Score 34; DB 4; Length 429;
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
 |||||
 Db 285 FLDKLNHNEI 293

RESULT 6
AAG81322
ID AAG81322 standard; protein; 429 AA.

XX
AC AAG81322;

XX
DT 10-SEP-2001 (first entry)

XX
DE Human AFP protein sequence SEQ ID NO:162.

XX
KW Human; secreted protein; secretion; bacterial cell; fungal cell;
KW eukaryotic cell; fusion protein; maltose binding protein;
KW immunoglobulin constant region; polyhistidine tag.

XX
OS Homo sapiens.

XX
PN WO200129221-A2.

XX
PD 26-APR-2001.

XX
PF 20-OCT-2000; 2000WO-US029052.

XX
PR 20-OCT-1999; 99US-0160712P.

XX
PA (ZYMO) ZYMOGENETICS INC.

XX
PI Conklin DC, Yee DP;

XX
DR WPI; 2001-300340/31.

XX
DR N-PSDB; AAH52173.

XX
PT Isolated polypeptide for directing secretion of proteins of interest from
PT a host cell including, e.g. bacteria, includes contiguous amino acid
PT residues of polypeptide with specified amino acids.

XX
PS Claim 1; Page 301-302; 617pp; English.

XX
CC AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242
CC to AAG81453. The secreted proteins can be used for directing the
CC secretion of proteins of interest from a host cell including bacteria,
CC fungal cells, and cultured higher eukaryotic cells. The present invention
CC also describes fusion proteins, where a secreted protein of the invention
CC is operably linked via a peptide bond or peptide linker to a second
CC protein selected from the group consisting of maltose binding protein, an
CC immunoglobulin constant region, a polyhistidine tag and a peptide given
CC in AAG81453

XX
SQ Sequence 429 AA;

Query Match 72.3%; Score 34; DB 4; Length 429;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9

DB 285 FLDKLNHNEI 293

RESULT 7
AAU81499
ID AAU81499 standard; protein; 429 AA.

XX
AC AAU81499;

XX
DT 09-APR-2002 (first entry)

XX
DE Mouse interflagellar transport protein, IFT57.

XX
KW Human; chlamydomonas; intraflagellar transport protein; Che-2;
KW antibacterial; protozoacide; contraceptive; antiinflammatory; malaria;
KW Chlamydomonas; IFT; mouse; kidney disease; retinal disorder; nematode;
KW thyroid disorder; chondrocyte disease; olfactory disease; azoospermia;

KW primary ciliary dyskinesia; insect; protozoa; male contraceptive;
KW parasitic disease; African sleeping sickness; trypanosomiasis;
KW leishmaniasis; trichomonosis; giardiasis; flagella.

XX
OS Chlamydomonas sp.

XX
PN WO200190307-A2.

XX
PD 29-NOV-2001.

XX
PF 24-MAY-2001; 2001WO-US017103.

XX
PR 24-MAY-2000; 2000US-0206923P.

XX
PA (UYMA-) UNIV MASSACHUSETTS.

XX
PI Witman GB, Pazour GJ, Rosenbaum JL, Cole DG;

XX
DR WPI; 2002-089926/12.

XX
DR N-PSDB; ABK27555.

XX
PT Novel Chlamydomonas or mouse intraflagellar transport (IFT) protein or
PT its variant, useful for identifying modulators that are useful for
PT treating nematode, insect, protozoa or bacterial infection by inhibiting
PT IFT.

XX
PS Claim 7; Fig 10D; 132pp; English.

XX
CC The invention relates to an isolated polypeptide (I) of Chlamydomonas
CC intraflagellar transport (IFT) particle protein, Che-2, or mouse IFT
CC particle protein 57. Also described are methods which are useful for:
CC diagnosing kidney disease, retinal disorder, thyroid disorder,
CC chondrocyte disease, olfactory disease, azoospermia, or primary ciliary
CC dyskinesia; for treating an infection in a mammal (a human) or plant
CC caused by nematode, insect, protozoa or bacteria. The nucleic acid
CC molecules are useful for diagnosis of disorders associated with aberrant
CC expression of nucleic acid molecules and in genetic mapping and
CC chromosome identification. The polypeptides are useful in generation of
CC antibodies, as reagents in diagnostic assays, for the identification of
CC other cellular gene products or compounds that can modulate the activity
CC or expression of nucleic acids or polypeptides, and as pharmaceutical
CC reagents useful for the treatment of disorders associated with aberrant
CC expression or activity of the nucleic acids or polypeptides. IFT can
CC serve as a male contraceptive. The IFT proteins can be targeted for
CC treating parasitic diseases such as malaria, African sleeping sickness,
CC trypanosomiasis, leishmanioses, trichomonosis, and giardiasis. An anti-IFT
CC drug, taken orally, would inhibit assembly of flagella in newly divided
CC Giardia and cause disassembly of previously formed flagella in non-
CC dividing Giardia. An anti-IFT drug would block assembly of the flagellum
CC and the flagellar sheath and affect the trypanosome's life cycle. The
CC anti-IFT modulators can also be used to treat trichomonos vaginitis in
CC humans. IFT inhibitors are also useful for combating phytopathogenic
CC nematodes inhibiting insect pest by blocking sensory cilia function
CC and/or assembly, leaving an insect unable to smell, unable to taste and
CC unable to hear. AAU81494-AAU81529 represent intraflagellar transport
CC particle amino acid sequences of the invention

XX
SQ Sequence 429 AA;

Query Match 72.3%; Score 34; DB 5; Length 429;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9

DB 285 FLDKLNHNEI 293

RESULT 8
AAU81512
ID AAU81512 standard; protein; 429 AA.

XX
AC AAU81512;

XX 09-APR-2002 (first entry)
DT Human interflagellar transport protein, IFT57 #1.
XX
XX
DE Human; chlamydomonas; intraflagellar transport protein; Che-2;
XX antibacterial; protozoacide; contraceptive; antiinflammatory; malaria;
KW Chlamydomonas; IFT; mouse; kidney disease; retinal disorder; nematode;
KW thyroid disorder; chondrocyte disease; olfactory disease; azoospermia;
KW primary ciliary dyskinesia; insect; protozoa; male contraceptive;
KW parasitic disease; African sleeping sickness; trypanosomiasis;
KW leishmaniasis; trichomonosis; giardiasis; flagella.
XX
OS Homo sapiens.
XX
XX WO200190307-A2.
PN 29-NOV-2001.
PD 24-MAY-2001; 2001WO-US017103.
XX 24-MAY-2000; 2000US-0206923P.
PR (UYMA-) UNIV MASSACHUSETTS.
XX Witman GB, Pazour GJ, Rosenbaum JL, Cole DG;
PI WPI; 2002-089926/12.
DR
XX Novel Chlamydomonas or mouse intraflagellar transport (IFT) protein or
PT its variant, useful for identifying modulators that are useful for
PT treating nematode, insect, protozoa or bacterial infection by inhibiting
PT IFT.
XX
PS Disclosure; Fig 10E; 132pp; English.
XX
CC The invention relates to an isolated polypeptide (I) of Chlamydomonas
CC intraflagellar transport (IFT) particle protein, Che-2, or mouse IFT
CC particle protein 57. Also described are methods which are useful for:
CC diagnosing kidney disease, retinal disorder, thyroid disorder,
CC chondrocyte disease, olfactory disease, azoospermia, or primary ciliary
CC dyskinesia; for treating an infection in a mammal (a human) or plant
CC caused by nematode, insect, protozoa or bacteria. The nucleic acid
CC molecules are useful for diagnosis of disorders associated with aberrant
CC expression of nucleic acid molecules and in genetic mapping and
CC chromosome identification. The polypeptides are useful in generation of
CC antibodies, as reagents in diagnostic assays, for the identification of
CC other cellular gene products or compounds that can modulate the activity
CC or expression of nucleic acids or polypeptides, and as pharmaceutical
CC reagents useful for the treatment of disorders associated with aberrant
CC expression or activity of the nucleic acids or polypeptides. IFT can
CC serve as a male contraceptive. The IFT proteins can be targeted for
CC treating parasitic diseases such as malaria, African sleeping sickness,
CC trypanosomiasis, leishmaniasis, trichomonosis, and giardiasis. An anti-IFT
CC drug, taken orally, would inhibit assembly of flagella in newly divided
CC Giardia and cause disassembly of previously formed flagella in non-
CC dividing Giardia. An anti-IFT drug would block assembly of the flagellum
CC and the flagellar sheath and affect the trypanosome's life cycle. The
CC anti-IFT modulators can also be used to treat trichomonos vaginitis in
CC humans. IFT inhibitors are also useful for combating phytopathogenic
CC nematodes inhibiting insect pest by blocking sensory cilia function
CC and/or assembly, leaving an insect unable to smell, unable to taste and
CC unable to hear. AAU81494-AAU81529 represent intraflagellar transport
CC particle amino acid sequences of the invention
XX
SQ Sequence 429 AA;

Query Match 72.3%; Score 34; DB 5; Length 429;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 FLDKAHMEV 9
| | | | | :
DB 285 FLDKLNHNEI 293

RESULT 10
ABG24971
ID ABG24971 standard; protein; 771 AA.
XX
AC ABG24971;
XX
DT 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #24962.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

Db 285 FLDKLNHNEI 293
RESULT 9
ADO48145
ID ADO48145 standard; protein; 429 AA.
XX
AC ADO48145;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human HIP-1.
XX
KW Huntingtin interacting protein 1; HIP-1; HIP-1 protein interactor;
KW apoptosis dysregulation.
XX
OS Homo sapiens.
XX
PN US2004096834-A1.
PD 20-MAY-2004.
XX
PF 19-NOV-2002; 2002US-00300263.
XX
PR 19-NOV-2002; 2002US-00300263.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Dobie KW;
XX
DR WPI; 2004-389149/36.
DR N-PSDB; ADO47996.
XX
PT New compounds targeted to a nucleic acid molecule encoding HIP-1 protein
PT interactor, useful for treating an animal having a disease or condition
PT associated with HIP-1 protein interactor, such as dysregulation of
PT apoptosis.
XX
PS Disclosure; Page 24-26; 76pp; English.
XX
CC The invention relates to a compound targeted to a nucleic acid molecule
CC encoding Huntingtin interacting protein 1 (HIP-1) protein interactor. The
CC compound is useful for treating an animal having a disease or condition
CC associated with HIP-1 protein interactor, such as dysregulation of
CC apoptosis. The compound may also be used for diagnostics, therapeutics,
CC prophylaxis and as research agents and kits; or to elucidate the function
CC of particular genes or to distinguish between functions of various
CC members of a biological pathway. The present sequence represents the
CC amino acid sequence of human HIP-1.
XX
SQ Sequence 429 AA;

Query Match 72.3%; Score 34; DB 8; Length 429;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 FLDKAHMEV 9
| | | | | :
DB 285 FLDKLNHNEI 293

RESULT 10
ABG24971
ID ABG24971 standard; protein; 771 AA.
XX
AC ABG24971;
XX
DT 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #24962.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.


```

XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS89158.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 55330; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activities. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 771 AA;

Query Match 72.3%; Score 34; DB 4; Length 771;
Best Local Similarity 62.5%; Pred. No. 3.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 122 FIDKGHLE 129

RESULT 11
ABG24799
ID ABG24799 standard; protein; 1683 AA.
XX AC ABG24799;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #24790.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

```

```

XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS88986.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 55158; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activities. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1683 AA;

Query Match 72.3%; Score 34; DB 4; Length 1683;
Best Local Similarity 62.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 674 FIDKGHLE 681

RESULT 12
ABR01783
ID ABR01783 standard; protein; 151 AA.
XX AC ABR01783;
XX DT 22-APR-2003 (first entry)
XX DE Human breast specific polypeptide #98.
XX KW Human; breast specific nucleic acid; BSNA; breast; cytostatic;
XX KW gene therapy; vaccines; lung cancer; breast cancer;
XX KW breast specific polypeptide; BSP.
XX OS Homo sapiens.

```


PT otitis media, conjunctivitis, pneumonia, bacteremia and especially
PT meningitis.
XX
PS Claim 11; Page 31-32; 40pp; English.
XX
CC The present sequence is the distal terminal portion of a spo/rel protein
CC (spoT/relA family) from Streptococcus pneumoniae. The protein is of the
CC spoT/relA family, which, in Escherichia coli, is involved in the
CC stringent response to nutrient limitation and regulate the accumulation
CC of (p)ppGpp which is involved in the regulation of gene expression and
CC other cellular processes. Spo/rel proteins have a homology to e.g. S.
CC equisimilis rel protein. Spo/rel proteins can be used: (a) in the
CC treatment of an individual in need of spo/rel protein, by administering
CC to the protein to the patient; (B) in the treatment of an individual
CC having need to inhibit spo/rel protein, by administering an antagonist
CC which inhibits the activity of the protein; and (C) for inducing an
CC immunological response by inoculating the mammal with spo/rel protein, or
CC a fragment or variant of it, adequate to produce antibody and/or T cell
CC immune response to protect the animal from disease. Conditions which may
CC be treated include otitis media, conjunctivitis, pneumonia, bacteremia,
CC meningitis, sinusitis, pleural empyema, endocarditis and especially
CC meningitis and compositions may be used as antibacterials.
XX
SQ Sequence 364 AA;

Query Match 70.2%; Score 33; DB 2; Length 364;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLDKAHMEV 9
|:|:|:|:
Db 135 FMDKGRHMD 142

RESULT 15
ABB98965
ID ABB98965 standard; protein; 366 AA.
XX
AC ABB98965;
XX
DT 14-APR-2003 (first entry)
XX
DE Human KIAA0615 protein 40.26.
XX
KW Human; KIAA0615 protein 40.26; dementia; facial paralysis; arrhythmia;
KW bronchial asthma; peptic ulcer; diabetes; periodical paralysis.
XX
OS Homo sapiens.
XX
PN CN1364786-A.
XX
PD 21-AUG-2002.
XX
PF 10-JAN-2001; 2001CN-00105159.
XX
PR 10-JAN-2001; 2001CN-00105159.
XX
PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y., Xie Y.
XX
WPI; 2003-000498/01.
DR N-PSDB; ABZ25074.
XX
PT New human KIAA0615 protein 40.26 polypeptide and encoding
PT polynucleotide, useful in treating e.g., diabetes and dementia.
XX
PS Claim 1; Page 27 (Disclosure); 34pp; Chinese.
XX
CC The present sequence is human KIAA0615 protein 40.26. The protein can be
CC used for treating various diseases, such as dementia, facial paralysis,
CC arrhythmia, bronchial asthma, peptic ulcer, diabetes and periodical
CC paralysis

XX SQ Sequence 366 AA;

Query Match 70.2%; Score 33; DB 6; Length 366;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLDKAHMEV 9
|:|:|:|:
Db 222 FLDKGHKDI 230

Search completed: December 14, 2004, 09:13:03
Job time : 10.9367 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:45:45 ; Search time 1.76856 Seconds
(without alignments)
337.485 Million cell updates/sec

Title: US-10-026-001-3
Perfect score: 51
Sequence: 1 FLDKGMVY 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	72.5	324	4	US-09-252-991A-22109
2	35	68.6	283	2	US-08-900-565-3
3	35	68.6	283	3	US-09-149-534-3
4	35	68.6	422	4	US-09-489-039A-12443
5	35	68.6	504	4	US-09-270-767-43244
6	35	68.6	652	4	US-09-252-991A-27811
7	34	66.7	997	4	US-10-101-464A-977
8	34	66.7	1002	4	US-10-101-464A-957
9	34	66.7	1203	4	US-09-661-258-3
10	34	66.7	1205	1	US-07-908-245-2
11	34	66.7	1205	2	US-08-319-866-10
12	34	66.7	1205	3	US-09-123-708-6
13	34	66.7	1205	3	US-09-123-624-6
14	34	66.7	1205	4	US-08-809-917-10
15	33	64.7	161	4	US-09-270-767-33785
16	33	64.7	161	4	US-09-270-767-49002
17	33	64.7	209	4	US-09-252-991A-23601
18	33	64.7	257	4	US-09-916-204-2
19	33	64.7	257	4	US-10-282-048-2
20	33	64.7	263	3	US-09-411-977-2
21	33	64.7	346	4	US-09-328-352-5835
22	33	64.7	362	4	US-09-489-039A-8250
23	33	64.7	497	4	US-09-804-471A-2
24	33	64.7	497	4	US-10-238-709-2
25	33	64.7	628	4	US-09-248-796A-18406
26	33	64.7	690	4	US-09-252-991A-27667
27	33	64.7	873	4	US-09-710-279-3036

28	33	64.7	876	1	US-08-785-071A-2	Sequence 2, Appli
29	33	64.7	876	3	US-09-012-872-2	Sequence 2, Appli
30	33	64.7	897	3	US-09-134-001C-3600	Sequence 3600, Ap
31	33	64.7	1702	3	US-08-296-791-5	Sequence 5, Appli
32	33	64.7	1702	4	US-09-839-996-5	Sequence 5, Appli
33	33	64.7	1702	4	US-10-080-505-5	Sequence 5, Appli
34	33	64.7	1702	5	PCT-US95-10661A-5	Sequence 5, Appli
35	33	64.7	1958	4	US-10-028-946-4	Sequence 4, Appli
36	33	64.7	2054	4	US-10-028-946-2	Sequence 2, Appli
37	32.5	63.7	899	3	US-09-413-814-5	Sequence 5, Appli
38	32	62.7	93	4	US-09-673-395A-417	Sequence 417, App
39	32	62.7	93	4	US-09-673-395A-462	Sequence 462, App
40	32	62.7	176	4	US-09-270-767-44425	Sequence 44425, A
41	32	62.7	295	4	US-09-613-303-33	Sequence 33, Appl
42	32	62.7	295	4	US-10-267-311-33	Sequence 33, Appl
43	32	62.7	312	4	US-09-543-681A-7100	Sequence 7100, Ap
44	32	62.7	324	4	US-09-489-039A-9892	Sequence 9892, Ap
45	32	62.7	334	4	US-09-252-991A-22708	Sequence 22708, A

ALIGNMENTS

RESULT 1
US-09-252-991A-22109
; Sequence 22109, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22109
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22109

Query Match 72.5%; Score 37; DB 4; Length 324;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGMVY 9
||| |||:
Db 163 FLDGGHIYM 171

RESULT 2
US-08-900-565-3
; Sequence 3, Application US/08900565
; Patent No. 5876996
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN S-ANDENOSYL-L-METHIONINE
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,565

FILING DATE: Herewith

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0352 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 283 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1065505

US-08-900-565-3

Query Match 68.6%; Score 35; DB 2; Length 283;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9

Db 72 LDAGHMFV 79

RESULT 3

US-09-149-534-3

Sequence 3, Application US/09149534

Patent No. 6379722

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

Lal, Preeti

Corley, Neil C.

Shah, Purvi

TITLE OF INVENTION: HUMAN S-ANDENOSYL-L-METHIONINE

METHYLTRANSFERASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/149,534

FILING DATE: 08-Sep-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/900,565

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0352 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 283 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1065505

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-149-534-3

Query Match 68.6%; Score 35; DB 3; Length 283;

Best Local Similarity 75.0%; Pred. No. 41;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9

Db 72 LDAGHMFV 79

RESULT 4

US-09-489-039A-12443

Sequence 12443, Application US/09489039A

Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 12443

LENGTH: 422

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-12443

Query Match 68.6%; Score 35; DB 4; Length 422;

Best Local Similarity 71.4%; Pred. No. 62;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKGHM 7

Db 87 FIDKGHL 93

RESULT 5

US-09-270-767-43244

Sequence 43244, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 43244

LENGTH: 504

TYPE: PRT

ORGANISM: Drosophila melanogaster

FEATURE:

OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43244

Query Match 68.6%; Score 35; DB 4; Length 504;
Best Local Similarity 62.5%; Pred. No. 74;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMY 8
| :|||:|
Db 78 FLRGHLY 85

RESULT 6

US-09-252-991A-27811
; Sequence 27811, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27811
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27811

Query Match 68.6%; Score 35; DB 4; Length 652;
Best Local Similarity 75.0%; Pred. No. 97;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9
| :|||:|
Db 464 LDGLHLYV 471

RESULT 7

US-10-101-464A-977
; Sequence 977, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 977
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-977

Query Match 66.7%; Score 34; DB 4; Length 997;

Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
| :|||:|
Db 511 FTEKGHIFV 519

RESULT 8

US-10-101-464A-957
; Sequence 957, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 957
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-957

Query Match 66.7%; Score 34; DB 4; Length 1002;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
| :|||:|
Db 536 FTEKGHIFV 544

RESULT 9

US-09-661-258-3
; Sequence 3, Application US/09661258
; Patent No. 6620616
; GENERAL INFORMATION:
; APPLICANT: Stuehr, Dennis J.
; APPLICANT: Adak, Subrata
; TITLE OF INVENTION: Nucleic Acids Encoding Nitric Oxide Synthase Variants
; FILE REFERENCE: 26473/04028
; CURRENT APPLICATION NUMBER: US/09/661,258
; CURRENT FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-661-258-3

Query Match 66.7%; Score 34; DB 4; Length 1203;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9
| :|||:|
Db 1106 LERGHMFV 1113

RESULT 10
US-07-908-245-2
; Sequence 2, Application US/07908245
; Patent No. 5498539
; GENERAL INFORMATION:
; APPLICANT: Harrison, David G.
; APPLICANT: Alexander, R. Wayne
; APPLICANT: Murphy, T.J.
; APPLICANT: Nishida, Ken'ichi
; TITLE OF INVENTION: Endothelial Nitric Oxide Synthase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,245
; FILING DATE: 19920702
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; TISSUE TYPE: Aorta
; CELL TYPE: Endothelial
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 496..512
; OTHER INFORMATION: /note= "CA++/CAM binding domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 651..678
; OTHER INFORMATION: /note= "FMN binding domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 795..806
; OTHER INFORMATION: /note= "FAD-Pyrophosphate binding
; OTHER INFORMATION: domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 937..947
; OTHER INFORMATION: /note= "FAD-Isolalloxanthine
; OTHER INFORMATION: binding domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1012..1030
; OTHER INFORMATION: /note= "NADPH-Ribose binding
; OTHER INFORMATION: domain"
; FEATURE:

; NAME/KEY: Binding-site
; LOCATION: 1111..1124
; OTHER INFORMATION: /note= "NADPH-Ribose binding
; OTHER INFORMATION: domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 33..34
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 46..47
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 53..54
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 58..59
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 97..98
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 116..117
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 282..283
; OTHER INFORMATION: /note= "Potential proline
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 459..460
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 472..473
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 602..603
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 727..728
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 838..839
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 869..870
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 872..873
; OTHER INFORMATION: /note= "Potential proline directed

OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 1085..1086
OTHER INFORMATION: /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 1202..1203
OTHER INFORMATION: /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 114..116
OTHER INFORMATION: /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 141..143
OTHER INFORMATION: /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 168..170
OTHER INFORMATION: /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 633..635
OTHER INFORMATION: /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 836..838
OTHER INFORMATION: /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 1051..1053
OTHER INFORMATION: /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION: phosphorylation site"
FEATURE: Domain
NAME/KEY: Domain
LOCATION: 738..740
OTHER INFORMATION: /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION: phosphorylation site"
US-07-908-245-2

Query Match 66.7%; Score 34; DB 1; Length 1205;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9
Db 1108 LERGHMFV 1115

RESULT 11
US-08-319-866-10
Sequence 10, Application US/08319866
Patent No. 5929223
GENERAL INFORMATION:
APPLICANT: Tully, Timothy P.
APPLICANT: Yin, Jerry C.
APPLICANT: Regulski, Michael
TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES
TITLE OF INVENTION: ASSOCIATED WITH LONG-TERM MEMORY
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts

COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/319,866
FILING DATE: 7-OCT-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: CSHL94-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1205 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-319-866-10

Query Match 66.7%; Score 34; DB 2; Length 1205;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9
Db 1108 LERGHMFV 1115

RESULT 12
US-09-123-708-6
Sequence 6, Application US/09123708
Patent No. 6146887
GENERAL INFORMATION:
APPLICANT: SCHRAEDER, Juergen
APPLICANT: GOEDECKE, Axel
TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
FILE REFERENCE: 511169-2003
CURRENT APPLICATION NUMBER: US/09/123,708
CURRENT FILING DATE: 1998-07-28
EARLIER APPLICATION NUMBER: 08/553,503
EARLIER FILING DATE: 1996-03-01
EARLIER APPLICATION NUMBER: P4411402.8
EARLIER FILING DATE: 1994-03-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 1205
TYPE: PRT
ORGANISM: Cytomegalovirus
US-09-123-708-6

Query Match 66.7%; Score 34; DB 3; Length 1205;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9
Db 1108 LERGHMFV 1115

RESULT 13

```
US-09-123-624-6
; Sequence 6, Application US/09123624
; Patent No. 6149936
; GENERAL INFORMATION:
; APPLICANT: SCHRAEDER, Jurgen
; APPLICANT: CODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2004
; CURRENT APPLICATION NUMBER: US/09/123,624
; CURRENT FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 08/553,503
; PRIOR FILING DATE: 1996-03-01
; PRIOR APPLICATION NUMBER: 4411402.8
; PRIOR FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1205
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-123-624-6

Query Match          66.7%; Score 34; DB 3; Length 1205;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LDKGHMYV 9
Db      1108 LERGHMFV 1115

RESULT 14
US-08-809-917-10
; Sequence 10, Application US/08809917
; Patent No. 6689557
; GENERAL INFORMATION:
; APPLICANT: APPLICANT
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES
; TITLE OF INVENTION: ASSOCIATED WITH LONG-TERM MEMORY
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,917
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13198
; FILING DATE:
; APPLICATION NUMBER: US 08/361,063
; FILING DATE: 21-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/319,866
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL94-03A2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
```

```
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-809-917-10

Query Match          66.7%; Score 34; DB 4; Length 1205;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LDKGHMYV 9
Db      1108 LERGHMFV 1115

RESULT 15
US-09-270-767-33785
; Sequence 33785, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33785
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33785

Query Match          64.7%; Score 33; DB 4; Length 161;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 LDKGHMY 8
Db      60 LDEGHIY 66

Search completed: December 14, 2004, 09:21:56
Job time : 2.76856 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 09:07:06 ; Search time 5.8559 Seconds
(without alignments)
548.952 Million cell updates/sec

Title: US-10-026-001-3

Perfect score: 51

Sequence: 1 FLDKGMVY 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	51	100.0	9	13	US-10-026-001-3
2	38	74.5	208	14	US-10-369-493-12076
3	38	74.5	979	15	US-10-424-599-258498
4	38	74.5	987	15	US-10-424-599-258499
5	37	72.5	59	15	US-10-424-599-285429
6	37	72.5	216	15	US-10-282-122A-65649
7	37	72.5	720	17	US-10-425-115-334973
8	37	72.5	729	15	US-10-264-049-2320
9	37	72.5	730	15	US-10-425-114-66927
10	37	72.5	852	14	US-10-097-340-57
11	37	72.5	852	14	US-10-177-293-76
12	37	72.5	1065	14	US-10-097-340-59
13	37	72.5	1065	14	US-10-177-293-78

14	37	72.5	1065	14	US-10-257-021-50	Sequence 50, Appl
15	37	72.5	1069	14	US-10-097-340-55	Sequence 55, Appl
16	37	72.5	1069	14	US-10-177-293-80	Sequence 80, Appl
17	37	72.5	1090	14	US-10-097-340-61	Sequence 61, Appl
18	36	70.6	9	13	US-10-026-001-5	Sequence 5, Appl
19	36	70.6	85	15	US-10-425-114-71617	Sequence 71617, A
20	36	70.6	150	15	US-10-424-599-144048	Sequence 144048, A
21	36	70.6	480	15	US-10-282-122A-67984	Sequence 67984, A
22	35	68.6	29	15	US-10-424-599-206447	Sequence 206447, A
23	35	68.6	40	15	US-10-424-599-174575	Sequence 174575, A
24	35	68.6	89	15	US-10-424-599-274631	Sequence 274631, A
25	35	68.6	94	15	US-10-424-599-204163	Sequence 204163, A
26	35	68.6	100	17	US-10-425-115-309889	Sequence 309889, A
27	35	68.6	233	16	US-10-437-963-125741	Sequence 125741, A
28	35	68.6	237	14	US-10-156-761-10738	Sequence 10738, A
29	35	68.6	242	15	US-10-424-599-211407	Sequence 211407, A
30	35	68.6	283	13	US-10-024-933-3	Sequence 3, Appl
31	35	68.6	339	15	US-10-424-599-148380	Sequence 148380, A
32	35	68.6	382	14	US-10-310-154-529	Sequence 529, App
33	35	68.6	419	17	US-10-425-115-209793	Sequence 209793, A
34	35	68.6	479	15	US-10-425-114-55580	Sequence 55580, A
35	35	68.6	479	17	US-10-425-115-316460	Sequence 316460, A
36	35	68.6	564	14	US-10-369-493-20223	Sequence 20223, A
37	35	68.6	566	15	US-10-425-114-46200	Sequence 46200, A
38	35	68.6	566	17	US-10-425-115-316462	Sequence 316462, A
39	34	66.7	86	17	US-10-425-115-251057	Sequence 251057, A
40	34	66.7	193	16	US-10-767-701-39702	Sequence 39702, A
41	34	66.7	289	13	US-10-135-322-10	Sequence 10, Appl
42	34	66.7	344	17	US-10-425-115-251318	Sequence 251318, A
43	34	66.7	550	14	US-10-369-493-18217	Sequence 18217, A
44	34	66.7	669	15	US-10-282-122A-48909	Sequence 48909, A
45	34	66.7	997	14	US-10-101-464A-977	Sequence 977, App

ALIGNMENTS

RESULT 1
US-10-026-001-3
; Sequence 3, Application US/10026001
; Publication No. US20020122791A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: ANTIGENIC CK-18 COMPOUNDS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS AND METHODS FOR USING SAME
; FILE REFERENCE: GZ 210800
; CURRENT APPLICATION NUMBER: US/10/026,001
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,820
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-001-3

Query Match 100.0%; Score 51; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKGMVY 9

Db 1 FLDKGMVY 9

RESULT 2

US-10-369-493-12076
; Sequence 12076, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12076
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12076

Query Match 74.5%; Score 38; DB 14; Length 208;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
||:||||
Db 77 FLERGHAYV 85

RESULT 3

US-10-424-599-258498
; Sequence 258498, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 258498
; LENGTH: 979
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75448C.1.pap
US-10-424-599-258498

Query Match 74.5%; Score 38; DB 15; Length 979;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
||:||||
Db 534 FTDKGHIFV 542

RESULT 4

US-10-424-599-258499
; Sequence 258499, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 258499
; LENGTH: 987
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75449C.1.pap
US-10-424-599-258499

Query Match 74.5%; Score 38; DB 15; Length 987;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
||:||||
Db 534 FTDKGHIFV 542

RESULT 5

US-10-424-599-285429
; Sequence 285429, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 285429
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_99771C.1.pap
US-10-424-599-285429

Query Match 72.5%; Score 37; DB 15; Length 59;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9
||:||||
Db 19 LEKGLYI 26

RESULT 6

US-10-282-122A-65649
; Sequence 65649, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlser, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65649
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-282-122A-65649

Query Match 72.5%; Score 37; DB 15; Length 216;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
|:|||||
Db 194 LEKGRMY 200

RESULT 7

US-10-425-115-334973
; Sequence 334973, Application US/10425115
; Publication No. US20040214272A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 334973
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_68609C.1.pap
US-10-425-115-334973

Query Match 72.5%; Score 37; DB 17; Length 720;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
|:|||||
Db 175 FVNGHAYV 183

RESULT 8

US-10-264-049-2320
; Sequence 2320, Application US/10264049
; Publication No. US20040005579A1

GENERAL INFORMATION:

; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2320
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2320

Query Match 72.5%; Score 37; DB 15; Length 729;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
|:|||||
Db 431 FLDKGEFYI 439

RESULT 9

US-10-425-114-66927
; Sequence 66927, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 66927
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17109D01_FLI.pap
US-10-425-114-66927

Query Match 72.5%; Score 37; DB 15; Length 730;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
|:|||||
Db 185 FVNGHAYV 193

RESULT 10

US-10-097-340-57
; Sequence 57, Application US/10097340
; Publication No. US20030087250A1

GENERAL INFORMATION:

; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT

```

; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-57

```

```

Query Match      72.5%; Score 37; DB 14; Length 852;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 FLDKGMVY 9
      |||||
Db      550 FLDKGEFYI 558

```

```

RESULT 11
US-10-177-293-76
; Sequence 76, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887

```

```

; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-76

```

```

Query Match      72.5%; Score 37; DB 14; Length 852;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 FLDKGMVY 9
      |||||
Db      550 FLDKGEFYI 558

```

```

RESULT 12
US-10-097-340-59
; Sequence 59, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 1065
; TYPE: PRT

```

```
; ORGANISM: Homo sapiens
US-10-097-340-59

Query Match      72.5%; Score 37; DB 14; Length 1065;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLDKGHMYV 9
      ||||| |
Db      767 FLDKGEFYI 775

RESULT 13
US-10-177-293-78
; Sequence 78, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Kamatkar, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-78

Query Match      72.5%; Score 37; DB 14; Length 1065;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLDKGHMYV 9
      ||||| |
Db      767 FLDKGEFYI 775

RESULT 14
US-10-257-021-50
; Sequence 50, Application US/10257021
; Publication No. US20030211498A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Morin, Patrice J.
; APPLICANT: Sherman-Baust, Cheryl A.
; APPLICANT: Pizer, Ellen S.
; APPLICANT: Hough, Colleen D.
; TITLE OF INVENTION: TUMOR MARKERS IN OVARIAN CANCER
; FILE REFERENCE: 14014.0369U2
; CURRENT APPLICATION NUMBER: US/10/257,021
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/10947
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/194,336
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-257-021-50

Query Match      72.5%; Score 37; DB 14; Length 1065;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      * 1 FLDKGHMYV 9
      ||||| |
Db      767 FLDKGEFYI 775

RESULT 15
US-10-097-340-55
; Sequence 55, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
```

; LENGTH: 1069
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-55

Query Match 72.5%; Score 37; DB 14; Length 1069;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
| | | | | | | |
Db 767 FLDKGFEYI 775

Search completed: December 14, 2004, 09:27:01
Job time : 6.8559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: December 14, 2004, 08:43:49 ; Search time 1.37555 Seconds
(without alignments)
629.533 Million cell updates/sec

Title: US-10-026-001-3
Perfect score: 51
Sequence: 1 FLDKGHMYV 9
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR 79: *
1: Pirl: *
2: Pirl2: *
3: Pirl3: *
4: Pirl4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	72.5	216	F82013	capsule polysaccha
2	37	72.5	216	H81241	capsule polysaccha
3	37	72.5	216	S15223	ctrD protein - Nei
4	37	72.5	239	G83143	hypothetical prote
5	37	72.5	470	C90398	cysteinyI-trNA syn
6	37	72.5	652	T03504	probable DNA topoi
7	37	72.5	1069	KUHU	ferroxidase (EC 1.
8	36	70.6	570	AH1927	hypothetical prote
9	36	70.6	967	SVMTAT	alanine-trNA ligas
10	35	68.6	259	F83825	hypothetical prote
11	35	68.6	283	T15659	hypothetical prote
12	35	68.6	320	T41927	hypothetical prote
13	35	68.6	344	T05437	hypothetical prote
14	35	68.6	348	H85256	hypothetical prote
15	35	68.6	407	AE0524	conserved hypothet
16	35	68.6	409	F90645	hypothetical prote
17	35	68.6	409	F85496	hypothetical prote
18	35	68.6	409	B64736	yadE protein - Esc
19	35	68.6	599	AB3375	beta-(1->2)glucan
20	35	68.6	1663	C3RT	complement C3 prec
21	34	66.7	368	A99270	sugar ABC transpor
22	34	66.7	459	T17644	hypothetical prote
23	34	66.7	532	G87912	protein B0205.9 [i
24	34	66.7	600	F84429	probable histidine
25	34	66.7	696	C85047	probable transposo
26	34	66.7	714	AF2479	ABC transporter Ar
27	34	66.7	926	T38198	COPII coated vesic
28	34	66.7	1203	A47501	nitric-oxide synth
29	34	66.7	1205	A38943	nitric-oxide synth

30	33	64.7	126	2	C70440	hypothetical prote
31	33	64.7	201	2	H83338	hypothetical prote
32	33	64.7	213	1	B71109	probable phosphoe
33	33	64.7	234	2	T37141	probable membrane
34	33	64.7	461	2	AC0569	cysteinyI-trNA syn
35	33	64.7	529	2	T20121	hypothetical prote
36	33	64.7	562	2	H70688	hypothetical prote
37	33	64.7	752	2	S73205	photosystem I prot
38	33	64.7	876	2	A89944	alanyl-trNA synthe
39	33	64.7	906	2	T01440	hypothetical prote
40	33	64.7	1202	2	S71424	nitric-oxide synth
41	33	64.7	1247	2	T31331	nitric-oxide synth
42	33	64.7	1694	2	H64106	IGA-specific metal
43	33	64.7	1702	2	A41859	IGA-specific metal
44	32	62.7	101	2	D82931	hypothetical prote
45	32	62.7	139	2	T26224	hypothetical prote

ALIGNMENTS

RESULT 1
F82013
capsule polysaccharide export ATP-binding protein NMA0195 [imported] - Neisseria meningi
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: F82013
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: F82013
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-216 <PAR>
A:Cross-references: UNIPROT:P57013; GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB8350
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: ctrD; NMA0195
C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 72.5%; Score 37; DB 2; Length 216;
Best Local Similarity 85.7%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
|:|||||
Db 194 LEKGHMY 200

RESULT 2
H81241
capsule polysaccharide export ATP-binding protein CtrD NMB0074 [imported] - Neisseria me
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81241
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, B.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: H81241
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-216 <TET>
A:Cross-references: UNIPROT:P32016; GB:AE002366; GB:AE002098; NID:g7225284; PIDN:AAF4054J
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0074
C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 72.5%; Score 37; DB 2; Length 216;
Best Local Similarity 85.7%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
|:|||||
Db 194 LDKGHMY 200

RESULT 3
S15223
ctrD protein - Neisseria meningitidis
C;Species: Neisseria meningitidis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S15223
R;Frosch, M.; Edwards, U.; Bousset, K.; Krause, B.; Weisgerber, C.
Mol. Microbiol. 5, 1251-1263, 1991
A;Title: Evidence for a common molecular origin of the capsule gene loci in Gram-negative
A;Reference number: S15220; MUID:92065822; PMID:1659649
A;Accession: S15223
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-216 <PRO>
A;Cross-references: UNIPROT:P32016; EMBL:M57677; NID:G150249; PIDN:AAA25453.1; PID:G1502
C;Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology
C;Keywords: ATP; nucleotide binding; P-loop
F;38-45/Region: nucleotide-binding motif A (P-loop)

Query Match 72.5%; Score 37; DB 2; Length 216;
Best Local Similarity 85.7%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
|:|||||
Db 194 LDKGHMY 200

RESULT 4
G83143
hypothetical protein PA4010 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: G83143
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83143
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-239 <STO>
A;Cross-references: UNIPROT:Q9HX17; GB:AE004818; GB:AE004091; NID:G9950200; PIDN:AAG0739
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4010

Query Match 72.5%; Score 37; DB 2; Length 239;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMY 9
|:|||||
Db 78 FLDKGHMY 86

RESULT 5
C90398
cysteinyI-tRNA synthetase (cyss) [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: C90398
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-y
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90398
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-470 <KUR>
A;Cross-references: UNIPROT:Q97WE6; GB:AE006641; NID:G13815580; PIDN:AAK42442.1; GSPDB:G
C;Genetics:
A;Gene: cyss
C;Superfamily: cysteine-tRNA ligase

Query Match 72.5%; Score 37; DB 2; Length 470;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKGHMY 9
|:|||||
Db 131 LDKGHMY 138

RESULT 6
T03504
probable DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain B - Rhodobacter capsulat
N;Alternate names: DNA gyrase chain b
C;Species: Rhodobacter capsulatus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03504
R;Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003.
A;Reference number: Z14955; MUID:97404404; PMID:9256491
A;Accession: T03504
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-652 <VLC>
A;Cross-references: UNIPROT:O68071; EMBL:AF010496; NID:G3128256; PIDN:AAC16157.1; PID:G31
C;Genetics:
A;Map position: 1
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B
C;Keywords: isomerase

Query Match 72.5%; Score 37; DB 2; Length 652;
Best Local Similarity 62.5%; Pred. No. 31;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 9
|:|||||
Db 536 LDKGHMY 543

RESULT 7
KURU
ferroxidase (EC 1.16.3.1) precursor [validated] - human
N;Alternate names: ceruloplasmin
N;Contains: ferroxidase long form (CP-1); ferroxidase short form (CP-2)
C;Species: Homo sapiens (man)
C;Date: 31-Aug-1980 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: A25443; A24165; A35450; A00524; I59067
R;Koschinsky, M.L.; Funk, W.D.; van Oost, B.A.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5086-5090, 1986
A;Title: Complete cDNA sequence of human preceruloplasmin.
A;Reference number: A25443; MUID:86259737; PMID:2873574
A;Accession: A25443
A;Molecule type: mRNA
A;Residues: 1-1060,1065-1069 <KOS>
A;Cross-references: UNIPROT:P00450; GB:M13699; NID:G180255; PIDN:AAA51976.1; PID:G180256
A;Note: this is the short or CP-2 alternatively spliced form

R;Mercer, J.F.B.; Grimes, A.
FEBS Lett. 203, 185-190, 1986
A;Title: Isolation of a human ceruloplasmin cDNA clone that includes the N-terminal lead
A;Reference number: A24165; MUID:86275241; PMID:3755405
A;Accession: A24165
A;Molecule type: mRNA
A;Residues: 1-40;549-599;784-829;919-952 <MER>
R;Yang, F.; Friedrichs, W.E.; Cupples, R.L.; Bonifacio, M.J.; Sanford, J.A.; Horton, W.A.
J. Biol. Chem. 265, 10780-10785, 1990
A;Title: Human ceruloplasmin. Tissue-specific expression of transcripts produced by alternative splicing
A;Reference number: A35450; MUID:90285218; PMID:2355023
A;Accession: A35450
A;Molecule type: DNA
A;Residues: 1007-1064 <YAN>
A;Cross-references: GB:J05506
A;Note: This is the long or CP-1 alternatively spliced form
R;Takahashi, N.; Ortel, T.L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 390-394, 1984
A;Title: Single-chain structure of human ceruloplasmin: the complete amino acid sequence
A;Reference number: A00524; MUID:84119493; PMID:6582496
A;Accession: A00524
A;Molecule type: protein
A;Residues: 20-1060,1065-1069 <TAK>
A;Note: 79-Gly and 449-Gly were also found
R;Yang, F.; Naylor, S.L.; Lum, J.B.; Cutshaw, S.; McCombs, J.L.; Naberhaus, K.H.; McGill
Proc. Natl. Acad. Sci. U.S.A. 83, 3257-3261, 1986
A;Title: Characterization, mapping, and expression of the human ceruloplasmin gene.
A;Reference number: I59067; MUID:86205876; PMID:3486416
A;Accession: I59067
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 218-1069 <RES>
A;Cross-references: GB:M13536; NID:G180248; PIDN:AAA51975.1; PID:G180249
C;Comment: Ferroxidase is a blue, plasma alpha2-glycoprotein binding 6-7 copper ions per molecule
C;Comment: In Wilson's disease the plasma levels of ferroxidase are diminished or undetectable
C;Comment: The three fragment chains are produced spontaneously during purification and
C;Genetics:
A;Gene: GDB:CP
A;Cross-references: GDB:I19069; OMIM:117700
A;Map position: 3q23-3q25
A;Introns: 1006/3; 1061/1
A;Note: the list of introns is incomplete
C;Function:
A;Description: catalyzes the oxidation of free iron(II) to iron(III) coupled with the reduction of NADPH to NADP+
A;Note: iron(III), but not iron(II), is the form bound and transported by transferrin
A;Note: other possible functions are amine oxidase activity, copper transport and homeostasis
C;Superfamily: ferroxidase; ferroxidase repeat homology
C;Keywords: acute phase; alternative splicing; copper; duplication; glycoprotein; oxidoreductase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1069/Product: ferroxidase, long form #status predicted <MATL>
F;20-1060/Product: ferroxidase, short form #status experimental <MARS>
F;20-499/Product: ferroxidase 67K chain #status experimental <K67>
F;23-357/Domain: ferroxidase repeat homology <FO1>
F;373-718/Domain: ferroxidase repeat homology <FO2>
F;501-905/Product: ferroxidase 50K chain #status experimental <K50>
F;733-1059/Domain: ferroxidase repeat homology <FO3>
F;907-1065/Product: ferroxidase 19K chain #status experimental <K19>
F;138,397,762/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;174-200,276-357,534-560,637-718,874-900/Disulfide bonds: #status predicted
F;227,588,926/Binding site: carbohydrate (Asn) (covalent) #status absent
F;295,338,343/Binding site: copper (His, Cys, His) (type 1) #status predicted
F;358/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;656,699,704,709/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted
F;994,1040,1045,1050/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 72.5%; Score 37; DB 1; Length 1069;
Best Local Similarity 66.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLDKGHMYV 9

Db 767 FLDKGEFYI 775

RESULT 8

AH1927
hypothetical protein alr0971 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH1927
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH1927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-570 <KUR>
A;Cross-references: UNIPROT:Q8YY80; GB:BA000019; PIDN:BA072928.1; PID:gl7130317; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0971
C;Superfamily: Synchocystis hypothetical protein slr1114

Query Match 70.6%; Score 36; DB 2; Length 570;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLDKGHMYV 9

Db 345 FLDKSHSYL 353

RESULT 9

SYNTAT
alanine-tRNA ligase (EC 6.1.1.7) - silkworm
C;Species: Bombyx mori (silkworm)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: A38327; S16072
R;Chang, P.K.; Dignam, J.D.
J. Biol. Chem. 265, 20898-20906, 1990
A;Title: Primary structure of alanyl-tRNA synthetase and the regulation of its mRNA level
A;Reference number: A38327; MUID:91065890; PMID:1701172
A;Accession: A38327
A;Molecule type: DNA
A;Residues: 1-967 <CHA>
A;Cross-references: UNIPROT:P21894; GB:M55993; GB:J055684; NID:gl55954; PIDN:AAA27821.1; I
R;Dignam, J.D.; Dignam, S.S.; Brumley, L.L.
Eur. J. Biochem. 198, 201-210, 1991
A;Title: Alanyl-tRNA synthetase from Escherichia coli, Bombyx mori and Rattus rattus. Exist
A;Reference number: S16072; MUID:91249799; PMID:2040280
A;Accession: S16072
A;Molecule type: DNA
A;Residues: 456-488 <DIG>
A;Cross-references: GB:J05684
C;Superfamily: alanyl-tRNA ligase
C;Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 70.6%; Score 36; DB 1; Length 967;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLDKGHMYV 9

Db 18 FIKKGHMYV 26

RESULT 10

F83825
hypothetical protein BH1406 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: F83825
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran

Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A;Reference number: A83650; MUID:20512582; PMID:11058132
 A;Accession: F83825
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-259 <STO>
 A;Cross-references: UNIPROT:Q9KD13; GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA051
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BH1406

Query Match 68.6%; Score 35; DB 2; Length 259;
 Best Local Similarity 85.7%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHM 7
 ||| |||
 Db 125 FLDNGHM 131

RESULT 11

T15659
 hypothetical protein C27F2.4 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C;Accession: T15659
 R;Kirsten, J.
 submitted to the EMBL Data Library, November 1995
 A;Description: The sequence of *C. elegans* cosmid C27F2.
 A;Reference number: Z18385
 A;Accession: T15659
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-283 <KIR>
 A;Cross-references: UNIPROT:Q18257; EMBL:U40419; NID:g1065504; PID:g1065505; PIDN:AAA814
 C;Genetics:
 A;Gene: CESP:C27F2.4
 A;Introns: 16/3; 61/1; 89/1; 123/2; 235/2

Query Match 68.6%; Score 35; DB 2; Length 283;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKGHMYV 9
 ||| |||
 Db 72 LDAGHMFV 79

RESULT 12

T41927
 hypothetical protein U25 - human herpesvirus 7 (strain JI)
 C;Species: human herpesvirus 7
 A;Variety: strain JI
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T41927
 R;Nicholas, J.
 submitted to the EMBL Data Library, December 1995
 A;Description: Determination and analysis of the complete nucleotide sequence of human h
 A;Reference number: Z22022
 A;Accession: T41927
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-320 <NIC>
 A;Cross-references: UNIPROT:P52528; EMBL:U43400; PIDN:AAC54687.1
 A;Experimental source: strain JI
 C;Genetics:
 A;Note: U25

Query Match 68.6%; Score 35; DB 2; Length 320;
 Best Local Similarity 71.4%; Pred. No. 35;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DKGHMYV 9
 ||| |||
 Db 91 DRGHLVY 97

RESULT 13

T05437
 hypothetical protein F7K2.10 - *Arabidopsis thaliana* (fragment)
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 C;Accession: T05437
 R;Bevan, M.; Medler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueler;
 submitted to the Protein Sequence Database, November 1998
 A;Reference number: Z15416
 A;Accession: T05437
 A;Molecule type: DNA
 A;Residues: 1-344 <BEV>
 A;Cross-references: EMBL:AL033545
 A;Experimental source: cultivar Columbia; BAC clone F7K2
 C;Genetics:
 A;Map position: 4
 A;Note: F7K2.10

Query Match 68.6%; Score 35; DB 2; Length 344;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
 ||| |||
 Db 101 FLNKGEMYM 109

RESULT 14

H85256
 hypothetical protein AT4G22430 [imported] - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: H85256
 R;anonymous; The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A;Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.
 A;Reference number: A85001; MUID:20083488; PMID:10617198
 A;Accession: H85256
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-348 <STO>
 A;Cross-references: UNIPROT:Q9SUX6; GB:NC_001268; NID:g7269089; PIDN:CA879198.1; GSPDB:G
 C;Genetics:
 A;Gene: AT4G22430
 A;Map position: 4

Query Match 68.6%; Score 35; DB 2; Length 348;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
 ||| |||
 Db 105 FLNKGEMYM 113

RESULT 15

AE0524
 conserved hypothetical protein STY0197 [imported] - *Salmonella enterica* subsp. *enterica* s
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A;Note: this species has also been called *Salmonella typhi*
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AE0524
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov.

A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AE0524
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-407 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01333.1; PID:gl6501461; GSPDB:GN00176
C;Genetics:
A;Gene: STY0197

Query Match 68.6%; Score 35; DB 2; Length 407;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FLDKGHM 7
|:|||||:
Db 72 FIDKGLH 78

Search completed: December 14, 2004, 09:20:20
Job time : 3.37555 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31 ; Search time 6.93668 Seconds
(without alignments)
465.433 Million cell updates/sec

Title: US-10-026-001-3
Perfect score: 51
Sequence: 1 FLDKGHMYV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	51	100.0	9	5 ABB81088	AB81088 Human cyt
2	37	72.5	216	6 ABU37725	ABU37725 Protein e
3	37	72.5	324	7 ABO73363	ABO73363 Pseudomon
4	37	72.5	729	5 ABP41188	ABP41188 Human ova
5	37	72.5	852	5 ABG96303	ABG96303 Human ova
6	37	72.5	852	6 ABR47422	ABR47422 Breast ca
7	37	72.5	1065	4 ABB50280	ABB50280 Cerulopla
8	37	72.5	1065	5 ABG96304	ABG96304 Human ova
9	37	72.5	1065	6 ABR47423	ABR47423 Breast ca
10	37	72.5	1065	7 ADD18608	ADD18608 Human dis
11	37	72.5	1065	7 ADE58878	ADE58878 Human Pro
12	37	72.5	1065	7 ADE57338	ADE57338 Human Pro
13	37	72.5	1065	7 ADD45526	ADD45526 Human Pro
14	37	72.5	1065	8 ADJ75542	ADJ75542 Marker ge
15	37	72.5	1069	5 ABG96302	ABG96302 Human ova
16	37	72.5	1069	5 ABR47424	ABR47424 Breast ca
17	37	72.5	1090	5 ABG96305	ABG96305 Human ova
18	36	70.6	9	5 ABB81089	ABB81089 Human cyt
19	36	70.6	480	6 ABU40060	ABU40060 Protein e
20	35	68.6	226	3 AAG48555	AAG48555 Arabidops
21	35	68.6	280	4 ABG02046	ABG02046 Novel hum
22	35	68.6	366	6 ABB98965	ABB98965 Human KIA
23	35	68.6	382	7 ADM48111	ADM48111 Polypepti
24	35	68.6	422	7 ABO65926	ABO65926 Klebsiell
25	35	68.6	459	8 ADQ07620	ADQ07620 Amino aci

26	35	68.6	467	4 ABB63044	Abb63044 Drosophil
27	35	68.6	652	7 ABO79065	ABO79065 Pseudomon
28	35	68.6	771	4 ABG24971	ABG24971 Novel hum
29	35	68.6	1663	8 ADI82104	ADI82104 Complemen
30	35	68.6	1683	4 ABG24799	ABG24799 Novel hum
31	34	66.7	107	4 AAB63525	AAB63525 Human gas
32	34	66.7	107	4 AAB63527	AAB63527 Human gas
33	34	66.7	107	4 AAB63524	AAB63524 Human gas
34	34	66.7	107	4 AAB63529	AAB63529 Human gas
35	34	66.7	289	5 AAO22563	AAO22563 Wooden le
36	34	66.7	337	4 ABB67824	ABB67824 Drosophil
37	34	66.7	600	5 ABB91688	ABB91688 Herbicida
38	34	66.7	630	4 ABB61655	ABB61655 Drosophil
39	34	66.7	669	6 ABU20985	ABU20985 Protein e
40	34	66.7	685	2 AAY25683	AAY25683 Cockroach
41	34	66.7	685	7 ADC34931	ADC34931 Cockroach
42	34	66.7	890	4 ABG13362	ABG13362 Novel hum
43	34	66.7	1036	5 ABG70782	ABG70782 A. thalia
44	34	66.7	1036	6 AAE33693	AAE33693 Arabidops
45	34	66.7	1057	5 AAO22558	AAO22558 Wooden le

ALIGNMENTS

RESULT 1
ABB81088
ID ABB81088 standard; peptide; 9 AA.
XX

AC ABB81088;
XX
DT 05-NOV-2002 (first entry)
XX

DE Human cytokeratin 18 (CK18) compound 1 peptide.

KW Antigen; cytokeratin 18; CK-18; immune response; cytostatic; vaccine;
KW gene therapy; cancer; human.
XX

OS Homo sapiens.

XX
PN WO200255555-A2.

XX
PD 18-JUL-2002.

XX
PF 21-DEC-2001; 2001WO-US049964.

XX
PR 21-DEC-2000; 2000US-0257820P.

XX
PA (GENZ) GENZYME CORP.

XX
PI Nicolette CA;

XX
DR WPI; 2002-619103/66.

XX
N-PSDB; ABN86611.

XX
PT Novel antigenic cytokeratin 18 compounds and peptides useful for inducing
PT an immune response in a subject and for diagnosing a neoplastic condition
or susceptibility to the condition of an animal cell or tissue.

XX
PS Claim 1; Page 63; 73pp; English.

XX
CC The invention relates to novel antigenic cytokeratin 18 (CK-18) compounds
(I) and peptides useful for inducing an immune response in a subject. (I)
is useful for inducing an immune response in a subject, by delivering (I)
as a polynucleotide, in the context of an MHC molecule that presents the
compound on the surface of an APC. (I) is useful for aiding in the
diagnosis of the neoplastic condition or susceptibility to the condition
of an animal cell or tissue; for generating antibodies which are useful
for identifying and purifying polypeptides and APCs expressing the
polypeptides. (I) serves as markers for the neoplastic phenotype. (I)
that is covalently or non-covalently linked to molecules are useful in
diagnostic methods, and for detecting or purifying antibodies. It is also
useful as components of anti-cancer vaccines and to expand immune

CC effector cells that are specific for cells having differential, i.e.
 CC aberrant, expression of antigenic CK-18. CK-18 proteins or antibodies are
 CC useful for detecting, diagnosing or prognosing and monitoring the
 CC progression, course or stage of CK-18 related cancers, or malignancies.
 CC Host cells comprising one or more immunogenic ligands are useful for
 CC inducing an immune response in a subject, and to expand a population of
 CC immune effector cells such as tumour infiltrating lymphocytes which in
 CC turn are useful in adoptive immunotherapies. Agents that modulate the
 CC binding of CK-18 protein to its ligand are useful for treating disease,
 CC especially cancer. The present sequence represents the human CK18
 CC compound 1 peptide
 XX Sequence 9 AA;
 SQ

Query Match 100.0%; Score 51; DB 5; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
 |||||
 Db 1 FLDKGHMYV 9

RESULT 2
 ABU37725
 ID ABU37725 standard; protein; 216 AA.
 XX
 AC ABU37725;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #23252.
 XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Neisseria meningitidis.
 XX

PN WO200277183-A2.
 XX

PD 03-OCT-2002.
 XX

PF 21-MAR-2002; 2002WO-US009107.
 XX

PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX

PA (ELIT-) ELITRA PHARM INC.
 XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 XX

DR WPI; 2003-029926/02.
 DR N-PSDB; ACA41595.
 XX

PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX

PS Claim 25; SEQ ID NO 65649; 1766pp; English.
 XX

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 216 AA;

Query Match 72.5%; Score 37; DB 6; Length 216;
 Best Local Similarity 85.7%; Pred. No. 26;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKGHMY 8
 |::|||
 Db 194 LDKGHMY 200

RESULT 3
 ABO73363
 ID ABO73363 standard; protein; 324 AA.
 XX

AC ABO73363;
 XX

DT 29-JUL-2004 (first entry)
 XX

DE Pseudomonas aeruginosa polypeptide #5538.
 XX

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX

OS Pseudomonas aeruginosa.
 OS

PN US6551795-B1.
 XX

PD 22-APR-2003.
 XX

PF 18-FEB-1999; 99US-00252991.
 XX

PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX

PA (GENO-) GENOME THERAPEUTICS CORP.
 XX

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX

DR WPI; 2003-615309/58.
 DR N-PSDB; ABD06934.
 XX

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX

PS Disclosure; SEQ ID NO 22109; 455pp; English.
 XX

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a

CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

XX Sequence 324 AA;

Query Match 72.5%; Score 37; DB 7; Length 324;
 Best Local Similarity 66.7%; Pred. No. 41;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9

Db 163 FLDGGHIYM 171

RESULT 4

ABP41188
 ID ABP41188 standard; protein; 729 AA.

AC ABP41188;

XX 22-AUG-2002 (first entry)

DT Human ovarian antigen HBLGD30, SEQ ID NO:2320.

DE Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 3q23-25.

XX Homo sapiens.

OS WO200200677-A1.

PN 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

PF 07-JUN-2000; 2000US-0209467P.

PR (HUMA-) HUMAN GENOME SCI INC.

XX Birse-CE, Rosen CA;

PI WPI; 2002-147878/19.

XX N-PSDB; ABQ54265.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.

XX Claim 11; SEQ ID NO 2320; 2922pp; English.

PS The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 729 AA;

Query Match 72.5%; Score 37; DB 5; Length 729;

Best Local Similarity 66.7%; Pred. No. 1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9

Db 431 FLDKGEFYI 439

RESULT 5

ABG96303

ID ABG96303 standard; protein; 852 AA.

XX ABG96303;

AC 11-DEC-2002 (first entry)

XX Human ovarian cancer marker OV7.

XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;
 KW central nervous system disorder; bacterial meningitis; viral meningitis;
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
 KW brain herniation; inflammation; encephalitis; testicular disorder;
 KW nontuberculous granulomatous orchitis; connective tissue disorder;
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
 KW histological type; carcinogenic; ovarian cancer marker.

XX Homo sapiens.

XX WO200271928-A2.

XX 19-SEP-2002.

XX 14-MAR-2002; 2002WO-US007826.

XX 14-MAR-2001; 2001US-0276025P.

PR 14-MAR-2001; 2001US-0276026P.

PR 10-AUG-2001; 2001US-0311732P.

PR 19-SEP-2001; 2001US-0323580P.

PR 26-SEP-2001; 2001US-0324967P.

PR 26-SEP-2001; 2001US-0325102P.

PR 26-SEP-2001; 2001US-0325149P.

XX (MILL-) MILLENNIUM PHARM INC.

PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
 PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
 PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
 XX WPI; 2002-723277/78.
 DR N-PSDB; ABS76395.
 XX
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient.
 XX
 PS Disclosure; Page 175-177; 481pp; English.
 XX
 CC The present invention relates to a new method for assessing whether a
 CC patient is afflicted with ovarian cancer. The method involves comparing
 CC the expression level of a marker in a patient sample and the normal level
 CC of expression of the marker in a control non-ovarian cancer sample, where
 CC the marker is selected from 363 cancer markers described in the
 CC specification. The method of the invention is useful in diagnosing or
 CC characterizing cancer, in detecting the presence of cancer as early as
 CC possible, and the recurrence of ovarian cancer. The method may also be of
 CC particular use with patients having an enhanced risk of developing
 CC ovarian cancer (e.g. patients having a familial history of ovarian
 CC cancer). The cancer markers may be used in the management and treatment
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
 CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
 CC disease or atherosclerosis). The compositions and methods may also be
 CC used in assessing the histological type of neoplasm associated with
 CC ovarian cancer, monitoring the progression of ovarian cancer, determining
 CC whether ovarian cancer has metastasized or is likely to metastasize,
 CC selecting a composition for inhibiting ovarian cancer, assessing the
 CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
 CC cancer or at risk of developing ovarian cancer. The present amino acid
 CC sequence represents one of the ovarian cancer markers described in the
 CC invention
 XX
 SQ Sequence 852 AA;
 Query Match 72.5%; Score 37; DB 5; Length 852;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 FLDKGHMYV 9
 |||||
 Db 550 FLDKGGEFYI 558
 RESULT 6
 ABR47422
 ID ABR47422 standard; protein; 852 AA.
 XX
 AC ABR47422;
 XX
 DT 12-JUN-2003 (first entry)
 XX
 DE Breast cancer associated protein sequence SEQ ID NO:76.
 XX
 KW Human; breast cancer; cytostatic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2003004989-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019669.
 XX
 PR 21-JUN-2001; 2001US-0299887P.

PR 27-JUN-2001; 2001US-0301572P.
 PR 18-JUL-2001; 2001US-0306501P.
 PR 25-SEP-2001; 2001US-0325002P.
 PR 05-MAR-2002; 2002US-0362585P.
 PR 14-MAY-2002; 2002US-0380391P.
 XX
 PA (MILL-) MILLENIUM PHARM INC.
 XX
 PI Lillie J, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S,
 PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
 PI Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
 XX
 DR WPI; 2003-210381/20.
 DR N-PSDB; ACC50114.
 XX
 PT Breast cancer diagnosis or treatment by comparing the level of expression
 PT of a marker in a patient sample with that in the control non-breast
 PT cancer sample.
 XX
 PS Claim 1; SEQ ID NO 76; 128pp; English.
 XX
 CC The present invention describes a method for assessing whether a patient
 CC is afflicted with breast cancer. The method comprises comparing the level
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
 CC ABR47386 to ABR47632) in a patient sample and the normal level of
 CC expression of the marker in a control non-breast cancer sample, where a
 CC significant increase in the level of expression of the marker in the
 CC patient sample and the normal level is an indication that the patient is
 CC afflicted with breast cancer. The breast cancer associated sequences from
 CC the present invention have cytostatic activities and can be used in gene
 CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 852 AA;
 Query Match 72.5%; Score 37; DB 6; Length 852;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 FLDKGHMYV 9
 |||||
 Db 550 FLDKGGEFYI 558
 RESULT 7
 ABB50280
 ID ABB50280 standard; protein; 1065 AA.
 XX
 AC ABB50280;
 XX
 DT 08-FEB-2002 (first entry)
 XX
 DE Ceruloplasmin (ferroxidase) ovarian tumour marker protein, SEQ ID NO:50.
 XX
 KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200175177-A2.
 XX
 PD 11-OCT-2001.


```
PF 03-APR-2001; 2001WO-US010947.
XX
PR 03-APR-2000; 2000US-0194336P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
XX
DR WPI; 2001-626450/72.
DR N-PSDB; ABA83106.
XX
XX Detecting and identifying ovarian tumor, identifying increased risk for
PT developing ovarian cancer, and determining effectiveness of ovarian
PT cancer treatment, by measuring expression level of ovarian tumor marker
PT gene.
XX
PS Claim 23; Page 94-96; 140pp; English.
XX
CC The invention relates to methods for diagnosing and prognosing ovarian
CC tumours in an individual via the detection and measurement of the
CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
CC ABA83181 and ABA83183). The methods of the invention are useful for
CC detecting an ovarian tumour in a patient, for identifying an individual
CC at increased risk for developing ovarian cancer, in prognostic tests for
CC assessing the relative severity of ovarian cancer, in tests for
CC monitoring a patient in remission from ovarian cancer and in tests for
CC monitoring disease status in a patient being treated for ovarian cancer.
CC The methods can additionally be used to identify a particular tumour as
CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
CC mucinous cystadenoma, borderline mucinous tumour, mucinous
CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
CC tumour. The ovarian tumour marker genes of the invention were identified
CC using SAGE (serial analysis of gene expression) and were found to be
CC overexpressed in a broad variety of ovarian epithelial tumour cells
CC relative to normal ovarian epithelial cells. The marker genes are
CC implicated in immune response pathways, in the regulation of cell
CC proliferation and in protein folding, and many of these are membrane-
CC localised or secreted. In addition to their use as diagnostic and
CC prognostic markers, the ovarian tumour marker genes or their encoded
CC proteins may be used as therapeutic targets for the treatment and
CC prevention of ovarian cancer. Sequences ABB50257-ABB50299 represent
CC proteins encoded by ovarian tumour marker genes of the invention
XX
SQ Sequence 1065 AA;

Query Match      72.5%; Score 37; DB 4; Length 1065;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGMVYV 9
Db 767 FLDKGEFYI 775

RESULT 8
ABG96304
ID ABG96304 standard; protein; 1065 AA.
XX
AC ABG96304;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker OV8.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;

histological type; carcinogenic; ovarian cancer marker.
Homo sapiens.
WO200271928-A2.
19-SEP-2002.
14-MAR-2002; 2002WO-US007826.
14-MAR-2001; 2001US-0276025P.
14-MAR-2001; 2001US-0276026P.
10-AUG-2001; 2001US-0311732P.
19-SEP-2001; 2001US-0323580P.
26-SEP-2001; 2001US-0324967P.
26-SEP-2001; 2001US-0325102P.
26-SEP-2001; 2001US-0325149P.
(MILL-) MILLENNIUM PHARM INC.
Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
WPI; 2002-723277/78.
N-PSDB; ABS76396.
Assessing whether a patient is afflicted with ovarian cancer, useful in
assessing the stage or progression of the disease, comprises comparing
the expression level of a cancer marker in a sample from a patient and
from a non cancer patient.
Disclosure; Page 178-180; 481pp; English.
The present invention relates to a new method for assessing whether a
patient is afflicted with ovarian cancer. The method involves comparing
the expression level of a marker in a patient sample and the normal level
of expression of the marker in a control non-ovarian cancer sample, where
the marker is selected from 363 cancer markers described in the
specification. The method of the invention is useful in diagnosing or
characterising cancer, in detecting the presence of cancer as early as
possible, and the recurrence of ovarian cancer. The method may also be of
particular use with patients having an enhanced risk of developing
ovarian cancer (e.g. patients having a familial history of ovarian
cancer). The cancer markers may be used in the management and treatment
of e.g. brain and central nervous system disorders (e.g. bacterial and
viral meningitis, Alzheimer's disease or Parkinson's disease), brain
disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
inflammations (e.g. bacterial or viral meningitis or encephalitis),
testicular disorders (e.g. nontuberculous granulomatous orchitis),
connective tissue disorders, or heart disorders (e.g. ischaemic heart
disease or atherosclerosis). The compositions and methods may also be
used in assessing the histological type of neoplasm associated with
ovarian cancer, monitoring the progression of ovarian cancer, determining
whether ovarian cancer has metastasized or is likely to metastasize,
selecting a composition for inhibiting ovarian cancer, assessing the
ovarian carcinogenic potential of a compound, or inhibiting ovarian
cancer or at risk of developing ovarian cancer. The present amino acid
sequence represents one of the ovarian cancer markers described in the
invention
XX
SQ Sequence 1065 AA;

Query Match      72.5%; Score 37; DB 5; Length 1065;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGMVYV 9
Db 767 FLDKGEFYI 775

RESULT 9
```


KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.

OS
XX WO2003016475-A2.
XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.
DR GENBANK; P00450.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1065 AA;

Query Match 72.5%; Score 37; DB 7; Length 1065;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
Db 767 FLDKGFEFYI 775

RESULT 12
ADE57338
ID ADE57338 standard; protein; 1065 AA.
XX
AC ADE57338;

XX 29-JAN-2004 (first entry)
DT
XX Human Protein P00450, SEQ ID NO 3199.

DE
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; P00450.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1065 AA;

Query Match 72.5%; Score 37; DB 7; Length 1065;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGHMYV 9
Db 767 FLDKGFEFYI 775

RESULT 13
ADD45526
ID ADD45526 standard; protein; 1065 AA.
XX
AC ADD45526;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P00450, SEQ ID NO 11190.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P00450.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1065 AA;
Query Match 72.5%; Score 37; DB 7; Length 1065;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX

QY 1 FLDKGMVY 9
Db 767 FLDKGMVY 775
XX
RESULT 14
ID ADJ75542 standard; protein; 1065 AA.
XX
AC ADJ75542;
XX
DT 20-MAY-2004 (first entry)
XX
DE Marker gene related amino acid sequence SEQ ID NO:794.
XX
KW bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.
XX
OS Homo sapiens.
XX
PN EP1394274-A2.
XX
PD 03-MAR-2004.
XX
PF 04-AUG-2003; 2003EP-00254857.
XX
PR 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX
PA (GENO-) GENOX RES INC.
XX
PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
XX WPI; 2004-193155/19.
DR
XX
PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX
PS Example 11; SEQ ID NO 794; 241pp; English.
XX
CC The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX

SQ Sequence 1065 AA;
Query Match 72.5%; Score 37; DB 8; Length 1065;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGMV 9
| | | | | |
Db 767 FLDKGEFYI 775

RESULT 15
ABG96302
ID ABG96302 standard; protein; 1069 AA.
XX
AC ABG96302;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker OV66.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
XX
PN WO200271928-A2.
XX
PD 19-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US007826.
XX
PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI Bast RC, Lu X, Schmandt RE, Zhao X, Glatt K;
XX
DR WPI; 2002-723277/78.
DR N-PSDB; ABS76394.
XX
PT Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
PS Disclosure; Page 172-174; 481pp; English.
XX
CC The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and

CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present amino acid
CC sequence represents one of the ovarian cancer markers described in the
CC invention
XX
SQ Sequence 1069 AA;
Query Match 72.5%; Score 37; DB 5; Length 1069;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKGMV 9
| | | | | |
Db 767 FLDKGEFYI 775

Search completed: December 14, 2004, 09:12:59
Job time : 9.93668 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:45:45 ; Search time 1.76856 Seconds
(without alignment)
337.485 Million cell updates/sec

Title: US-10-026-001-5

Perfect score: 47

Sequence: 1 FLDKAHMEV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:**

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	74.5	143	4 US-09-270-767-35197	Sequence 35197, A
2	35	74.5	143	4 US-09-270-767-50414	Sequence 50414, A
3	34	72.3	422	4 US-09-489-039A-12443	Sequence 12443, A
4	33	70.2	364	3 US-08-891-322-6	Sequence 6, Appli
5	33	70.2	364	3 US-09-277-019-6	Sequence 6, Appli
6	33	70.2	740	3 US-08-891-322-2	Sequence 2, Appli
7	33	70.2	740	3 US-09-277-019-2	Sequence 2, Appli
8	33	70.2	740	4 US-09-583-110-3225	Sequence 3225, Ap
9	32	68.1	143	4 US-09-252-991A-26374	Sequence 26374, A
10	32	68.1	212	4 US-09-543-681A-4702	Sequence 4702, Ap
11	32	68.1	336	4 US-09-252-991A-31353	Sequence 31353, A
12	32	68.1	481	1 US-07-754-918A-13	Sequence 13, Appli
13	32	68.1	494	4 US-09-198-452A-215	Sequence 215, App
14	32	68.1	566	4 US-09-431-470-2	Sequence 2, Appli
15	32	68.1	635	3 US-09-341-833A-6	Sequence 6, Appli
16	32	68.1	635	3 US-09-341-833A-7	Sequence 7, Appli
17	32	68.1	635	3 US-09-341-833A-9	Sequence 9, Appli
18	32	68.1	677	3 US-09-341-833A-8	Sequence 8, Appli
19	31	66.0	145	4 US-09-732-210-327	Sequence 327, App
20	31	66.0	360	4 US-09-252-991A-26045	Sequence 26045, A
21	31	66.0	630	3 US-09-342-647-2	Sequence 2, Appli
22	31	66.0	1199	3 US-09-208-742-2	Sequence 2, Appli
23	31	66.0	1199	3 US-09-332-295-4	Sequence 4, Appli
24	31	66.0	1199	4 US-09-709-979-4	Sequence 4, Appli
25	31	66.0	1199	4 US-10-147-268-4	Sequence 4, Appli
26	31	66.0	1495	4 US-08-522-726B-1	Sequence 1, Appli
27	31	66.0	1495	4 US-09-337-384-1	Sequence 1, Appli

28 30 63.8 87 4 US-09-107-532A-5006 Sequence 5006, Ap
29 30 63.8 149 4 US-09-270-767-32526 Sequence 32526, A
30 30 63.8 149 4 US-09-270-767-47743 Sequence 47743, A
31 30 63.8 159 4 US-09-248-796A-22395 Sequence 22395, A
32 30 63.8 180 4 US-09-248-796A-17490 Sequence 17490, A
33 30 63.8 260 4 US-09-270-767-37235 Sequence 37235, A
34 30 63.8 260 4 US-09-270-767-52452 Sequence 52452, A
35 30 63.8 346 1 US-07-661-610C-4 Sequence 4, Appli
36 29 61.7 43 4 US-09-149-476-355 Sequence 355, App
37 29 61.7 49 4 US-09-270-767-61731 Sequence 61731, A
38 29 61.7 133 4 US-09-543-681A-7988 Sequence 7988, Ap
39 29 61.7 136 3 US-09-134-001C-5296 Sequence 5296, Ap
40 29 61.7 138 4 US-09-134-000C-3646 Sequence 3646, Ap
41 29 61.7 176 4 US-09-270-767-44425 Sequence 44425, A
42 29 61.7 224 4 US-09-270-767-32918 Sequence 32918, A
43 29 61.7 224 4 US-09-270-767-48135 Sequence 48135, A
44 29 61.7 250 4 US-09-252-991A-30269 Sequence 30269, A
45 29 61.7 259 4 US-09-270-767-39034 Sequence 39034, A

ALIGNMENTS

RESULT 1
US-09-270-767-35197
; Sequence 35197, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35197
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35197

Query Match 74.5%; Score 35; DB 4; Length 143;
Best Local Similarity 55.6%; Pred. No. 8.3;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:|:|:|:
Db 39 FIDKQHIEI 47

RESULT 2
US-09-270-767-50414
; Sequence 50414, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50414
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50414

Query Match 74.5%; Score 35; DB 4; Length 143;

```
Best Local Similarity 55.6%; Pred. No. 8.3;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
Db 39 FIDKQHIEI 47

RESULT 3
US-09-489-039A-12443
; Sequence 12443, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12443
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12443

Query Match 72.3%; Score 34; DB 4; Length 422;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 87 FIDKQHLE 94

RESULT 4
US-08-891-322-6
; Sequence 6, Application US/08891322
; Patent No. 6096518
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; TITLE OF INVENTION: No. 6096518el Compounds
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,322
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/029049
; FILING DATE: 24-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50576
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-891-322-6

Query Match 70.2%; Score 33; DB 3; Length 364;
Best Local Similarity 62.5%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 135 FMDKRRHMD 142

RESULT 5
US-09-277-019-6
; Sequence 6, Application US/09277019
; Patent No. 6268179
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; TITLE OF INVENTION: No. 6268179el Compounds
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/277,019
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,322
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: 60/029049
; FILING DATE: 24-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50576
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-277-019-6

Query Match 70.2%; Score 33; DB 3; Length 364;
Best Local Similarity 62.5%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 135 FMDKRRHMD 142

RESULT 6
US-09-277-019-6
; Sequence 6, Application US/08891322
; Patent No. 6096518
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; TITLE OF INVENTION: No. 6096518el Compounds
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,322
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/029049
; FILING DATE: 24-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50576
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-891-322-6

Query Match 70.2%; Score 33; DB 3; Length 364;
Best Local Similarity 62.5%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 135 FMDKRRHMD 142
```

US-08-891-322-2
; Sequence 2, Application US/08891322
; Patent No. 6096518
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; TITLE OF INVENTION: No. 6096518el Compounds
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,322
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/029049
; FILING DATE: 24-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50576
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 740 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-891-322-2

Query Match 70.2%; Score 33; DB 3; Length 740;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 511 FMDKRHMD 518

RESULT 7
US-09-277-019-2
; Sequence 2, Application US/09277019
; Patent No. 6268179
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; TITLE OF INVENTION: No. 6268179el Compounds
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

US-08-891-322-2
; APPLICATION NUMBER: US/09/277,019
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,322
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: 60/029049
; FILING DATE: 24-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50576
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 740 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-277-019-2

Query Match 70.2%; Score 33; DB 3; Length 740;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 511 FMDKRHMD 518

RESULT 8
US-09-583-110-3225
; Sequence 3225, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3225
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3225

Query Match 70.2%; Score 33; DB 4; Length 740;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
Db 511 FMDKRHMD 518

RESULT 9
US-09-252-991A-26374
; Sequence 26374, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 26374
;; LENGTH: 143
;; TYPE: PRT

;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26374

Query Match 68.1%; Score 32; DB 4; Length 143;
Best Local Similarity 77.8%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLDKAHMEV 9
||| ||| |||
Db 121 FLDLAHDEV 129

RESULT 10

US-09-543-681A-4702
; Sequence 4702, Application US/09543681A
; Patent No. 6505709

;; GENERAL INFORMATION:

;; APPLICANT: GARY BRETON

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

;; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

;; FILE REFERENCE: 2709.1002-001

;; CURRENT APPLICATION NUMBER: US/09/543,681A

;; CURRENT FILING DATE: 2000-04-05

;; PRIOR APPLICATION NUMBER: US 60/128,706

;; PRIOR FILING DATE: 1999-04-09

;; NUMBER OF SEQ ID NOS: 8344

;; SEQ ID NO 4702

;; LENGTH: 212

;; TYPE: PRT

;; ORGANISM: Proteus mirabilis

US-09-543-681A-4702

Query Match 68.1%; Score 32; DB 4; Length 212;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLDKAHME 8
||| ||| |||
Db 105 FLTKAHVE 112

RESULT 11

US-09-252-991A-31353
; Sequence 31353, Application US/09252991A
; Patent No. 6551795

;; GENERAL INFORMATION:

;; APPLICANT: Marc J. Rubenfield et al.

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

;; FILE REFERENCE: 107196.136

;; CURRENT APPLICATION NUMBER: US/09/252,991A

;; CURRENT FILING DATE: 1999-02-18

;; PRIOR APPLICATION NUMBER: US 60/074,788

;; PRIOR FILING DATE: 1998-02-18

;; PRIOR APPLICATION NUMBER: US 60/094,190

;; PRIOR FILING DATE: 1998-07-27

;; NUMBER OF SEQ ID NOS: 33142

;; SEQ ID NO 31353

;; LENGTH: 336

;; TYPE: PRT

;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31353

Query Match 68.1%; Score 32; DB 4; Length 336;
Best Local Similarity 62.5%; Pred. No. 86;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDKAHMEV 9
||| ||| |||
Db 260 IDQAHLEV 267

RESULT 12

US-07-754-918A-13

; Sequence 13, Application US/07754918A

; Patent No. 5286484

;; GENERAL INFORMATION:

;; APPLICANT: Rodriguez, R.S. et al

;; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR AN

;; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN FROM NEISSERIA MENINGITIDIS AND USE

;; TITLE OF INVENTION: OF SAID PROTEIN IN VACCINE PREPARATIONS

;; NUMBER OF SEQUENCES: 14

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Stanger, Michaelson, Spivak and Wallace, Esq.

;; STREET: Parkway 109 Office Center, 328 Newman Springs

;; CITY: Red Bank

;; STATE: New Jersey

;; COUNTRY: USA

;; ZIP: 07701

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: 5 1/4" 360Kb IBM compatible diskette

;; COMPUTER: IBM PS/2 Model 80

;; OPERATING SYSTEM: MS-DOS 5.0

;; SOFTWARE: Microsoft Word 5.5

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/07/754,918A

;; FILING DATE: 19910905

;; CLASSIFICATION: 424

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Michaelson, Peter L.

;; REGISTRATION NUMBER: 30090

;; REFERENCE/DOCKET NUMBER: Centro-2R

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (908)530-6671

;; TELEFAX: (908)530-6584

;; INFORMATION FOR SEQ ID NO: 13:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 481 amino acids

;; TYPE: AMINO ACID

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: P64k fragment comprising the homology region

;; MOLECULE TYPE: with Lipamide Dehydrogenase from E. coli

US-07-754-918A-13

Query Match 68.1%; Score 32; DB 1; Length 481;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLDKAHMEV 9
||| ||| |||
Db 117 FLDPHLEV 125

RESULT 13

US-09-198-452A-215

; Sequence 215, Application US/09198452A

; Patent No. 6559294

;; GENERAL INFORMATION:

;; APPLICANT: Griffais, R.

;; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention

; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 215
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-215

Query Match 68.1%; Score 32; DB 4; Length 494;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKAHMEV 9
||| |:
Db 119 LDKRHLEI 126

RESULT 14

US-09-431-470-2
; Sequence 2, Application US/09431470
; Patent No. 6433249
; GENERAL INFORMATION:
; APPLICANT: Duvick, Jon
; APPLICANT: Simmons, Carl R.
; APPLICANT: Crasta, Oswald R.
; APPLICANT: Folkerts, Otto
; TITLE OF INVENTION: The Use of Beta-Glucosidase to Enhance
; FILE REFERENCE: 5718-43
; CURRENT APPLICATION NUMBER: US/09/431,470
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,920
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Zea mays
US-09-431-470-2

Query Match 68.1%; Score 32; DB 4; Length 566;
Best Local Similarity 55.6%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
||| |:
Db 210 FLDKSHKSI 218

RESULT 15

US-09-341-833A-6
; Sequence 6, Application US/09341833A
; Patent No. 6383488
; GENERAL INFORMATION:
; APPLICANT: CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA
; TITLE OF INVENTION: PRE-M/M PROTEIN EPITOPES OF THE DENGUE VIRUS, SYNTHETIC
; FILE REFERENCE: Dengue prem/M SEQs 1-9 for 976-4(0003)
; CURRENT APPLICATION NUMBER: US/09/341,833A
; CURRENT FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: CU 13/97
; PRIOR FILING DATE: 1997-01-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 6
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
; OTHER INFORMATION: Of Dengue virus prem/M and Neisseria meningitidis
; OTHER INFORMATION: P64k protein.
US-09-341-833A-6

Query Match 68.1%; Score 32; DB 3; Length 635;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
||| |:
Db 271 FLDPHLEV 279

Search completed: December 14, 2004, 09:21:57
Job time : 2.76856 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004. Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 09:07:06 ; Search time 5.8559 Seconds
(without alignments)
548.952 Million cell updates/sec

Title: US-10-026-001-5

Perfect score: 47

Sequence: 1 FLDKARMEV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PT_NEW PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	9	13	US-10-026-001-5
2	36	76.6	9	13	US-10-026-001-3
3	34	72.3	65	17	US-10-425-115-200286
4	34	72.3	386	16	US-10-437-963-122283
5	34	72.3	429	9	US-09-866-582-12
6	34	72.3	429	9	US-09-866-582-30
7	34	72.3	504	17	US-10-411-910A-78
8	34	72.3	887	16	US-10-437-963-122285
9	33	70.2	63	17	US-10-425-115-274014
10	33	70.2	86	17	US-10-425-115-318154
11	33	70.2	151	13	US-10-001-843-213
12	33	70.2	186	16	US-10-437-963-159258
13	33	70.2	740	9	US-09-815-242-13483

14	33	70.2	740	15	US-10-282-122A-74108	Sequence 74108, A
15	33	70.2	813	16	US-10-437-963-159271	Sequence 159271, A
16	33	70.2	971	14	US-10-369-493-5225	Sequence 5225, Ap
17	33	70.2	1039	14	US-10-369-493-2257	Sequence 2257, Ap
18	33	70.2	1039	14	US-10-369-493-2258	Sequence 2258, Ap
19	32	68.1	121	14	US-10-369-493-12660	Sequence 12660, A
20	32	68.1	128	17	US-10-425-115-192793	Sequence 192793, A
21	32	68.1	138	17	US-10-425-115-205799	Sequence 205799, A
22	32	68.1	190	15	US-10-282-122A-54315	Sequence 54315, A
23	32	68.1	198	15	US-10-282-122A-69070	Sequence 69070, A
24	32	68.1	200	14	US-10-369-493-12702	Sequence 12702, A
25	32	68.1	220	17	US-10-425-115-347351	Sequence 347351, A
26	32	68.1	252	17	US-10-425-115-331536	Sequence 331536, A
27	32	68.1	253	17	US-10-425-115-195027	Sequence 195027, A
28	32	68.1	256	17	US-10-425-115-308551	Sequence 308551, A
29	32	68.1	319	15	US-10-425-114-66685	Sequence 66685, A
30	32	68.1	326	15	US-10-282-122A-51774	Sequence 51774, A
31	32	68.1	348	15	US-10-424-599-166656	Sequence 166656, A
32	32	68.1	390	15	US-10-425-114-38490	Sequence 38490, A
33	32	68.1	394	14	US-10-369-493-8385	Sequence 8385, Ap
34	32	68.1	463	17	US-10-425-115-331535	Sequence 331535, A
35	32	68.1	464	15	US-10-425-114-69690	Sequence 69690, A
36	32	68.1	474	15	US-10-425-114-70966	Sequence 70966, A
37	32	68.1	474	17	US-10-425-115-331539	Sequence 331539, A
38	32	68.1	494	15	US-10-289-762-215	Sequence 215, App
39	32	68.1	507	15	US-10-282-122A-48984	Sequence 48984, A
40	32	68.1	508	15	US-10-425-114-46495	Sequence 46495, A
41	32	68.1	510	14	US-10-369-493-4244	Sequence 4244, Ap
42	32	68.1	527	14	US-10-022-832-6	Sequence 6, Appli
43	32	68.1	527	15	US-10-312-273-127	Sequence 127, App
44	32	68.1	566	17	US-10-425-115-331542	Sequence 331542, A
45	32	68.1	567	15	US-10-425-114-65837	Sequence 65837, A

ALIGNMENTS

RESULT 1
US-10-026-001-5
; Sequence 5, Application US/10026001
; Publication No. US20020122791A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: ANTIGENIC CK-18 COMPOUNDS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS AND METHODS FOR USING SAME
; FILE REFERENCE: GZ 210800
; CURRENT APPLICATION NUMBER: US/10/026,001
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,820
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-001-5

Query Match 100.0%; Score 47; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKARMEV 9

Db 1 FLDKARMEV 9

RESULT 2

US-10-026-001-3
; Sequence 3, Application US/10026001
; Publication No. US20020122791A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.

; TITLE OF INVENTION: ANTIGENIC CK-18 COMPOUNDS FOR THERAPY
; FILE REFERENCE: G2 210800
; CURRENT APPLICATION NUMBER: US/10/026,001
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,820
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-001-3

Query Match 76.6%; Score 36; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
Db 1 FLDKGHMYV 9

RESULT 3

US-10-425-115-200286
; Sequence 200286, Application US/10425115
; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 200286

; LENGTH: 65

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_114248C.1.pep

US-10-425-115-200286

Query Match 72.3%; Score 34; DB 17; Length 65;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDKAHMEV 9
Db 13 LDKAHVDV 20

RESULT 4

US-10-437-963-122283
; Sequence 122283, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122283
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_25225C.1.pep
US-10-437-963-122283

Query Match 72.3%; Score 34; DB 16; Length 386;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DKAHMEV 9
Db 355 DKAHLEV 361

RESULT 5

US-09-866-582-12

; Sequence 12, Application US/09866582

; Patent No. US20020127620A1

; GENERAL INFORMATION:

; APPLICANT: Witman, George B.

; APPLICANT: Pazour, Gregory J.

; APPLICANT: Rosenbaum, Joel L.

; APPLICANT: Cole, Douglas G.

; TITLE OF INVENTION: INTRACELLULAR TRANSPORT

; FILE REFERENCE: 07917-145001

; CURRENT APPLICATION NUMBER: US/09/866,582

; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 60/206,923

; PRIOR FILING DATE: 2000-05-24

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 429

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-866-582-12

Query Match 72.3%; Score 34; DB 9; Length 429;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
Db 285 FLDKLHNEI 293

RESULT 6

US-09-866-582-30

; Sequence 30, Application US/09866582

; Patent No. US20020127620A1

; GENERAL INFORMATION:

; APPLICANT: Witman, George B.

; APPLICANT: Pazour, Gregory J.

; APPLICANT: Rosenbaum, Joel L.

; APPLICANT: Cole, Douglas G.

; TITLE OF INVENTION: INTRACELLULAR TRANSPORT

; FILE REFERENCE: 07917-145001

; CURRENT APPLICATION NUMBER: US/09/866,582

; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 60/206,923

; PRIOR FILING DATE: 2000-05-24

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 30

; LENGTH: 429

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-866-582-30

Query Match 72.3%; Score 34; DB 9; Length 429;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
| | | | | | | | | |
Db 285 FLDKLHNEI 293

RESULT 7
US-10-411-910A-78
; Sequence 78, Application US/10411910A
; Publication No. US20040209256A1
; GENERAL INFORMATION:
; APPLICANT: Dillion, Harrison F.
; TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes
; FILE REFERENCE: H2041203-P
; CURRENT APPLICATION NUMBER: US/10/411,910A
; CURRENT FILING DATE: 2003-04-12
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 78
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Entamoeba histolytica
US-10-411-910A-78

Query Match 72.3%; Score 34; DB 17; Length 504;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
| | | | | | | | | |
Db 385 FVDKAHKE 392

RESULT 8
US-10-437-963-122285
; Sequence 122285, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122285
; LENGTH: 887
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_25227C.1.pap
US-10-437-963-122285

Query Match 72.3%; Score 34; DB 16; Length 887;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DKAHMEV 9
| | | | | | | | | |
Db 442 DKAHLEV 448

RESULT 9

US-10-425-115-274014
; Sequence 274014, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274014
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(63)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_181488C.1.pap
US-10-425-115-274014

Query Match 70.2%; Score 33; DB 17; Length 63;
Best Local Similarity 62.5%; Pred. No. 54;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
| | | | | | | | | |
Db 24 YLDRAHIE 31

RESULT 10

US-10-425-115-318154
; Sequence 318154, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 318154
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_53233C.1.pap
US-10-425-115-318154

Query Match 70.2%; Score 33; DB 17; Length 86;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKAHMEV 9
| | | | | | | | | |
Db 40 LDKAHPEI 47

RESULT 11

US-10-001-843-213
; Sequence 213, Application US/10001843
; Publication No. US20020132255A1
; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and PR
; FILE REFERENCE: DEX-0267
; CURRENT APPLICATION NUMBER: US/10/001,843
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,992
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 213
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-843-213

Query Match 70.2%; Score 33; DB 13; Length 151;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:|||||
Db 78 FVDKAHME 85

RESULT 12
US-10-437-963-159258
; Sequence 159258, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159258
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58652C.1.pap
US-10-437-963-159258

Query Match 70.2%; Score 33; DB 16; Length 186;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:|||||
Db 3 FLNPAHLEV 11

RESULT 13
US-09-815-242-13483
; Sequence 13483, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13483
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13483

Query Match 70.2%; Score 33; DB 9; Length 740;
Best Local Similarity 62.5%; Pred. No. 7.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:|||||
Db 511 FMDKRRMD 518

RESULT 14
US-10-282-122A-74108
; Sequence 74108, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74108
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-74108

Query Match 70.2%; Score 33; DB 15; Length 740;
Best Local Similarity 62.5%; Pred. No. 7.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:|:|:|:
Db 511 FMDKRHMD 518

RESULT 15
US-10-437-963-159271
; Sequence 159271, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159271
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58664C.1.pep
US-10-437-963-159271

Query Match 70.2%; Score 33; DB 16; Length 813;
Best Local Similarity 55.6%; Pred. No. 8.2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:|:|:|:
Db 448 FMDKLHMAI 456

Search completed: December 14, 2004, 09:27:02
Job time : 6.8559 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:49 ; Search time 1.37555 Seconds
(without alignments)
629.533 Million cell updates/sec

Title: US-10-026-001-5
Perfect score: 47
Sequence: 1 FLDKAHMEV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	35	74.5	594	D55514	dihydrolipoamide d
2	35	74.5	682	G71299	conserved hypothe
3	34	72.3	409	F90645	hypothetical prote
4	34	72.3	409	F85496	hypothetical prote
5	34	72.3	409	B64736	yadE protein - Esc
6	34	72.3	621	T48187	hypothetical prote
7	33	70.2	157	A10978	probable membrane
8	33	70.2	157	S47798	hypothetical 17.5K
9	33	70.2	181	C69847	ribosomal-protein-
10	33	70.2	396	H87709	HlyD family secret
11	33	70.2	740	D95191	GTP pyrophosphokin
12	33	70.2	740	F98057	GTP diphosphokinas
13	33	70.2	971	T32883	hypothetical prote
14	33	70.2	1039	T43678	tetratricopeptide
15	33	70.2	1039	T38447	tetratricopeptide
16	32	68.1	190	B81386	probable periplasm
17	32	68.1	215	T36637	probable DNA-bind
18	32	68.1	229	T21921	hypothetical prote
19	32	68.1	268	A11578	cobalt transport A
20	32	68.1	326	A97316	hypothetical prote
21	32	68.1	338	T24933	hypothetical prote
22	32	68.1	341	T24927	hypothetical prote
23	32	68.1	417	T19381	hypothetical prote
24	32	68.1	440	E71625	variant-specific s
25	32	68.1	448	T32710	hypothetical prote
26	32	68.1	527	B72107	peptide ABC transp
27	32	68.1	527	D86515	oligopeptide bindi
28	32	68.1	566	A48860	beta-glucosidase,
29	32	68.1	594	G81847	dihydrolipoamide d

30	32	68.1	594	1	S42920	dihydrolipoamide d
31	32	68.1	645	2	S49570	penicillin-binding
32	31	66.0	135	2	B40655	probable replicati
33	31	66.0	145	2	H69082	ribosomal protein
34	31	66.0	226	2	C71289	hypothetical prote
35	31	66.0	259	2	F83825	hypothetical prote
36	31	66.0	314	2	T45673	hypothetical prote
37	31	66.0	317	2	T19143	hypothetical prote
38	31	66.0	325	2	T19451	hypothetical prote
39	31	66.0	359	2	G70860	probable fecB - My
40	31	66.0	443	2	B70933	hypothetical prote
41	31	66.0	446	2	B87912	protein B0205.6 [i
42	31	66.0	558	2	T48078	hypothetical prote
43	31	66.0	571	2	B64469	hypothetical prote
44	31	66.0	597	2	F84512	Mutator-like trans
45	31	66.0	629	2	T48799	Rrp9p related prot

ALIGNMENTS

RESULT 1

D55514
C:Species: Alcaligenes eutrophus
C:Date: 18-Aug-1995 #sequence_revision 10-May-1996 #text_change 09-Jul-2004
C:Accession: D55514
R:Hein, S.; Steinbuechel, A.
J. Bacteriol. 176, 4394-4408, 1994
A:Title: Biochemical and molecular characterization of the Alcaligenes eutrophus pyruvat
A:Reference number: A55514; MUID:94292470; PMID:8021225
A:Accession: D55514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-594 <HEI>
A:Cross-references: UNIPROT:Q59099; GB:U09865; NID:G497263; PIDN:AAA21600.1; PID:G497266
C:Genetics:
A:Gene: pdhL
C:Superfamily: Alkaligenes dihydrolipoamide dehydrogenase; dihydrolipoamide dehydrogenase
C:Keywords: FAD; flavoprotein; lipoamide; NAD; oxidoreductase; redox-active disulfide
F:5-77/Domain: lipoyl/biotin-binding homology <LPB>
F:122-150/Region: beta-alpha-beta FAD nucleotide-binding fold
F:124-574/Domain: dihydrolipoamide dehydrogenase homology <DLD>
F:298-326/Region: beta-alpha-beta NAD nucleotide-binding fold
F:43/Binding site: lipoamide (Lys) (covalent) #status predicted
F:159-164/Disulfide bonds: redox-active #status predicted

Query Match 74.5%; Score 35; DB 1; Length 594;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLDKAHMEV 9
Db 233 FLDPHMEV 241

RESULT 2

G71299
conserved hypothetical protein TP0648 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: G71299
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
son, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
n, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: G71299
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-682 <COL>
A:Cross-references: UNIPROT:O83654; GB:AE001239; GB:AE000520; NID:G3322937; PIDN:AAC6562J

A;Experimental source: strain Nichols

C;Genetics:

A;Gene: TP0648

C;Superfamily: hypothetical protein TP0648

Query Match 74.5%; Score 35; DB 1; Length 682;
Best Local Similarity 55.6%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9

Db 437 YLDKAHLSI 445

RESULT 3

F90645 hypothetical protein ECs0134 [imported] - Escherichia coli (strain O157:H7, substrain R1)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: F90645

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90645

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-409 <HAY>

A;Cross-references: UNIPROT:Q8X933; GB:BA000007; PIDN:BA033557.1; PID:gl13359590; GSPDB:C

A;Experimental source: strain O157:H7, substrain R1MD 0509952

C;Genetics:

A;Gene: ECs0134

Query Match 72.3%; Score 34; DB 2; Length 409;

Best Local Similarity 62.5%; Pred. No. 34;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8

Db 74 FIDKGHLE 81

RESULT 4

F85496 hypothetical protein yadE [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: F85496

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamocousis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: F85496

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-409 <STO>

A;Cross-references: UNIPROT:Q8X933; GB:AE005174; NID:gl2512845; PIDN:AAG54434.1; GSPDB:C

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: yadE

Query Match 72.3%; Score 34; DB 2; Length 409;

Best Local Similarity 62.5%; Pred. No. 34;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8

Db 74 FIDKGHLE 81

RESULT 5

B64736

yadE protein - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: B64736

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B64736

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-409 <BLAT>

A;Cross-references: UNIPROT:P31666; GB:AE000122; GB:U00096; NID:gl1786315; PIDN:AAC73241.1

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: yadE

C;Keywords: transmembrane protein

F;5-21/Domain: transmembrane #status predicted <TMM>

Query Match 72.3%; Score 34; DB 2; Length 409;

Best Local Similarity 62.5%; Pred. No. 34;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8

Db 74 FIDKGHLE 81

RESULT 6

T48187

hypothetical protein F7A7.180 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C;Accession: T48187

R;Bevan, M.; Terry, N.; Ardiles, W.; Buysschaert, C.; Dasseville, R.; De Clerck, R.; De

ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24487

A;Accession: T48187

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-621 <BEV>

A;Cross-references: UNIPROT:Q9M010; EMBL:AL161946

A;Experimental source: cultivar Columbia; BAC clone F7A7

C;Genetics:

A;Map position: 5

A;Introns: 65/3; 263/3; 289/3; 403/1; 469/3; 495/2; 529/2; 592/3

A;Note: F7A7.180

Query Match 72.3%; Score 34; DB 2; Length 621;

Best Local Similarity 75.0%; Pred. No. 54;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDKAHMEV 9

Db 289 LDKAHKEI 296

RESULT 7

AI0978

probable membrane protein STY4127 [imported] - Salmonella enterica subsp. enterica serov

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AI0978

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0978
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-157 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD07957.1; PID:g16504945; GSPDB:GN00176
C;Genetics:
A;Gene: STY4127

Query Match 70.2%; Score 33; DB 2; Length 157;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:| |:|
Db 57 FMDNAHVQV 65

RESULT 8

S47798
hypotheical 17.5K protein (avtA-selB intergenic region) - Escherichia coli (strain K-12)
N;Alternate names: hypothetical protein o157a
C;Species: Escherichia coli

C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S47798; C65157

R;Plunkett, G.
submitted to the EMBL Data Library, March 1994

A;Reference number: S47798
A;Accession: S47798

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-157 <PLU>

A;Cross-references: UNIPROT:P37674; EMBL:U00039; NID:g466582; PIDN:AA18554.1; PID:g4667

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: C65157

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-157 <BLAT>

A;Cross-references: GB:AF000435; GB:U00096; NID:g2367244; PIDN:AAC76601.1; PID:g1790002;

A;Experimental source: strain K-12, substrain MGL655

C;Genetics:

A;Gene: yiam

Query Match 70.2%; Score 33; DB 2; Length 157;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:| |:|
Db 57 FMDNAHVQV 65

RESULT 9

C69847
ribosomal-protein-alanine N-acetyltransfer homolog yjck - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: C69847

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togroni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: C69847

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-181 <KUN>

A;Cross-references: UNIPROT:O31633; GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13046.

A;Experimental source: strain 168

C;Genetics:

A;Gene: yjck

C;Superfamily: Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimJ

Query Match 70.2%; Score 33; DB 1; Length 181;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKAH 6
|:| |:|
Db 99 FLDKAH 104

RESULT 10

H87709

HlyD family secretion protein [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: H87709

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: H87709

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-396 <STO>

A;Cross-references: UNIPROT:Q9A252; GB:AE005673; NID:gl3425482; PIDN:AAK25676.1; GSPDB:G

C;Genetics:

A;Gene: CC3714

C;Superfamily: multidrug resistance protein A; lipoyl/biotin-binding homology

Query Match 70.2%; Score 33; DB 2; Length 396;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:| |:|
Db 59 FLDKQHF 66

RESULT 11

D95191

GTP pyrophosphokinase [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C;Accession: D95191

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I

son, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: D95191

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-740 <KUR>

A;Cross-references: UNIPROT:Q97PH2; GB:AE005672; PIDN:AAK75725.1; PID:g14973135; GSPDB:G

A;Experimental source: strain TIGR4

C;Genetics:
A;Gene: SP1645
C;Superfamily: guanosine 3',5'-bis(diphosphate) 3'-pyrophosphatase

Query Match 70.2%; Score 33; DB 2; Length 740;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:|:|:|:
Db 511 FMDKRRHMD 518

RESULT 12
F98057
GTP diphosphokinase (EC 2.7.6.5) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: F98057
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: F98057
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-740 <KUR>
A;Cross-references: UNIPROT:Q8DNX1; GB:AB007317; PIDN:AAL00291.1; PID:G15459146; GSPDB:G

C;Genetics:
A;Gene: rela
C;Superfamily: guanosine 3',5'-bis(diphosphate) 3'-pyrophosphatase
C;Keywords: diphosphotransferase

Query Match 70.2%; Score 33; DB 2; Length 740;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:|:|:|:
Db 511 FMDKRRHMD 518

RESULT 13
T32883
hypothetical protein C34B2.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32883
R;Graves, T.; Sutterer, C.; Hawkins, M.; Wilson, R.
submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid C34B2.
A;Reference number: Z21241
A;Accession: T32883
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-971 <GRA>
A;Cross-references: UNIPROT:O44952; EMBL:AF043693; PIDN:AAB97538.1; GSPDB:GN00019; CESP:
A;Experimental source: strain Bristol N2; clone, C34B2
C;Genetics:
A;Map position: 1
A;Introns: 74/3; 108/3; 139/2; 289/2; 343/3; 448/2; 532/3; 622/2; 708/3; 788/3; 869/3; 9
C;Superfamily: ATP-dependent Lon protease

Query Match 70.2%; Score 33; DB 2; Length 971;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:|:|:|:
Db 465 FLESAMSV 473

Search completed: December 14, 2004, 09:20:21
Job time: 2.37555 secs

Db 857 FPDKAHIHI 865

RESULT 14

T43678
tetrairicopeptide repeat protein tp1 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T43678
R;Lichtenberg-Frate, H.; Heyer, M.; Hoefler, M.
submitted to the EMBL Data Library, February 1998
A;Description: Tp1, a novel Schizosaccharomyces pombe protein with several tetrairico p
A;Reference number: Z22625
A;Accession: T43678
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-1039 <LIC>
A;Cross-references: EMBL:AF047464; PIDN:AA03120.1
A;Experimental source: strain 972
C;Genetics:
A;Gene: tp1

Query Match 70.2%; Score 33; DB 2; Length 1039;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:|:|:|:
Db 465 FLESAMSV 473

RESULT 15

T38447
tetrairicopeptide repeat protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38447; T38994
R;McDougall, R.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1997
A;Reference number: Z21793
A;Accession: T38447
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1039 <MCD>
A;Cross-references: UNIPROT:O42668; EMBL:AL009227; PIDN:CAA15833.1; GSPDB:GN00066; SPDB:G
A;Experimental source: strain 972h; cosmid c27D7
R;Purnelle, B.; Goffeau, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21823
A;Accession: T38994
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1039 <PUR>
A;Cross-references: EMBL:AL034583; PIDN:CAA22581.1; GSPDB:GN00066; SPDB:SPAC637.02C
A;Experimental source: strain 972h; cosmid c637
C;Genetics:
A;Gene: SPAC27D7.14c; SPAC637.02c
A;Map position: 1

Query Match 70.2%; Score 33; DB 2; Length 1039;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
|:|:|:|:
Db 465 FLESAMSV 473

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31 ; Search time 7.11354 Seconds
(without alignments)
727.960 Million cell updates/sec

Title: US-10-026-001-5
Perfect score: 47
Sequence: 1 FLDKAHMEV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	80.9	312	2 Q6NLSO	Q6nls0 arabidopsis
2	38	80.9	312	2 AAS49045	Aas49045 arabidops
3	38	80.9	312	2 AAS76747	Aas76747 arabidops
4	36	76.6	497	2 Q6RIA6	Q6ria6 abudeiduf s
5	36	76.6	497	2 AAR89626	Aar89626 abudeiduf
6	35	74.5	234	2 Q96DA9	Q96da9 homo sapien
7	35	74.5	594	2 Q59099	Q59099 alcaligenes
8	35	74.5	682	2 O83654	O83654 treponema p
9	34	72.3	172	2 Q8KJQ3	Q8kjq3 escherichia
10	34	72.3	174	2 Q7WS17	Q7ws17 alcaligenes
11	34	72.3	220	2 Q9Y0D9	Q9y0d9 strongyloce
12	34	72.3	272	2 Q9CUS6	Q9cus6 mus musculu
13	34	72.3	275	2 Q6VCY8	Q6vcy8 cowdria rum
14	34	72.3	275	2 AAR10933	Aar10933 cowdria r
15	34	72.3	342	1 ASTE VIBPA	Q87q40 vibrio para
16	34	72.3	384	2 Q83ME7	Q83me7 shigella fl
17	34	72.3	409	1 YADE ECOLI	P31666 escherichia
18	34	72.3	409	2 Q7AHF0	Q7ahp0 escherichia
19	34	72.3	409	2 Q7UDR5	Q7udr5 shigella fl
20	34	72.3	409	2 Q8CWE1	Q8cwe1 escherichia
21	34	72.3	409	2 Q82RR2	Q8zrr2 salmonella
22	34	72.3	409	2 Q8X933	Q8x933 escherichia
23	34	72.3	429	2 Q9NWS7	Q9nwb7 homo sapien
24	34	72.3	429	2 Q8BXG3	Q8bxg3 mus musculu
25	34	72.3	429	2 Q924M2	Q924m2 mus musculu
26	34	72.3	429	2 CAG33532	Cag33532 homo sapi
27	34	72.3	440	2 BAC28265	Bac28265 m 16 days
28	34	72.3	504	2 Q869B1	Q869b1 entamoeba h
29	34	72.3	515	2 Q815R1	Q815r1 bacillus ce
30	34	72.3	569	2 Q84H24	Q84h24 pseudomonas
31	34	72.3	594	2 Q6DCR8	Q6dcr8 xenopus lae

32	34	72.3	621	2 QSM010	Qsm010 arabidopsis
33	34	72.3	658	2 Q73NB1	Q73nb1 treponema d
34	34	72.3	658	2 AAS11762	Aas11762 treponema
35	34	72.3	1020	2 Q6MEX9	Q6mex9 parachlamyd
36	34	72.3	1020	2 CAP22870	Cap22870 parachlam
37	34	72.3	1581	2 Q6C494	Q6c494 yarrowia li
38	33	70.2	135	2 Q6K7B6	Q6k7b6 oryza sativ
39	33	70.2	135	2 BAD19566	Bad19566 oryza sat
40	33	70.2	157	1 YIAM ECOLI	P37674 escherichia
41	33	70.2	157	2 Q8XG27	Q8xg27 salmonella
42	33	70.2	157	2 Q7CPI2	Q7cpi2 salmonella
43	33	70.2	172	1 LOLA WOLSU	O7mrk6 wolinnella s
44	33	70.2	181	2 O31633	O31633 bacillus su
45	33	70.2	188	1 FX36_MOUSE	Q9cqc24 mus musculu

ALIGNMENTS

RESULT 1

Q6NLSO PRELIMINARY; PRT; 312 AA.
AC Q6NLSO;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE At2g29070.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT012260; AAS76747.1; -
DR EMBL; BT011682; AAS49045.1; -
DR InterPro; IPR004854; UFD1.
DR Pfam; PF03152; UFD1; 1.
SQ SEQUENCE 312 AA; 34902 MW; A4197813955654D4 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 312;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 FLDKAHME 8
DB 19 FIDKAHLE 26
:::|||||::

RESULT 2

AAS49045 PRELIMINARY; PRT; 312 AA.
AC AAS49045;
DT 10-MAR-2004 (Tremblrel. 27, Created)
DT 10-MAR-2004 (Tremblrel. 27, Last sequence update)
DT 10-MAR-2004 (Tremblrel. 27, Last annotation update)
DE At2g29070.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT011682; AAS49045.1; -
SQ SEQUENCE 312 AA; 34902 MW; A4197813955654D4 CRC64;
Query Match 80.9%; Score 38; DB 2; Length 312;

Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:|||||:
Db 19 FIDKAHLE 26

RESULT 3

AAS76747 PRELIMINARY; PRT; 312 AA.
AC AAS76747;
DT 29-MAR-2004 (TREMBlrel. 27, Created)
DT 29-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 29-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE A2929070.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT012260; AAS76747.1; -;
SQ SEQUENCE 312 AA; 34902 MW; A4197813955654D4 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 312;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLDKAHME 8
|:|||||:
Db 19 FIDKAHLE 26

RESULT 4

Q6RIA6 PRELIMINARY; PRT; 497 AA.
AC Q6RIA6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE 14-alpha demethylase.
GN Name=CYP51;
OS Abudufduf saxatilis (Sergeant major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidel;
OC Pomacentridae; Abudufduf.
OX NCBI_TaxID=50731;
RN [1]
RP SEQUENCE FROM N.A.
RA Morrison A.M., Stegeman J.J.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY498940; AAR89626.1; -;
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PRO0465; EP450IV.
DR PRINTS; PRO0385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Methyltransferase.
SQ SEQUENCE 497 AA; 56607 MW; 3DD5E1EC24A313B9 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 497;
Best Local Similarity 85.7%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DKAHMEV 9
|:|||||:
Db 248 DKAHMEI 254

RESULT 5

AAR89626 PRELIMINARY; PRT; 497 AA.
AC AAR89626;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE 14-alpha demethylase.
GN CYP51.
OS Abudufduf saxatilis (Sergeant major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidel;
OC Pomacentridae; Abudufduf.
OX NCBI_TaxID=50731;
RN [1]
RP SEQUENCE FROM N.A.
RA Morrison A.M., Stegeman J.J.;
RT "Molecular phylogeny of CYP51";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY498940; AAR89626.1; -;
KW Methyltransferase.
SQ SEQUENCE 497 AA; 56607 MW; 3DD5E1EC24A313B9 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 497;
Best Local Similarity 85.7%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DKAHMEV 9
|:|||||:
Db 248 DKAHMEI 254

RESULT 6

Q96DA9 PRELIMINARY; PRT; 234 AA.
AC Q96DA9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ESRRL1 protein (Fragment).
GN Name=ESRRL1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009690; AAH09690.2; -.
FT NON TER 1
SQ SEQUENCE 234 AA; 27069 MW; AC56D7A36D51A264 CRC64;

Query Match 74.5%; Score 35; DB 2; Length 234;
Best Local Similarity 77.8%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
||| |
Db 90 FLDKLHNEV 98

RESULT 7
Q59099 PRELIMINARY; PRT; 594 AA.
AC Q59099;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dihydrolipoamide dehydrogenase (EC 1.8.1.4).
GN Name=pdhL;
OS Alkaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Wautersia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H16;
RX MEDLINE=9432470; PubMed=8021225;
RA Hein S., Steinbuechel A.;
RT "Biochemical and molecular characterization of the Alkaligenes
RT eutrophus pyruvate dehydrogenase complex and identification of a new
RT type of dihydrolipoamide dehydrogenase.";
RL J. Bacteriol. 176:4394-4408 (1994).
CC -|- CATALYTIC ACTIVITY: Protein N(6)-(dihydrolipoyl)lysine + NAD(+) =
CC protein N(6)-(lipoyl)lysine + NADH.
CC -|- COFACTOR: Binds 1 FAD per subunit (By similarity).
CC -|- COFACTOR: FAD (By similarity).
CC -|- MISCELLANEOUS: The active site is a redox-active disulfide bond
CC (By similarity).
CC -|- SIMILARITY: Belongs to the class-I pyridine nucleotide-disulfide
CC oxidoreductase family.
CC -|- SIMILARITY: Contains 1 lipoyl-binding domain.
DR EMBL; U09865; AAA21600.1; -.
DR PIR; D55514; D55514.
DR HSSP; Q51225; 1OUT.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004148; F:dihydrolipoyl dehydrogenase activity; IEA.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR000815; Hg_reductase.
DR InterPro; IPR011053; Hybrid_motif.
DR InterPro; IPR006258; Lipoamide_dh.
DR InterPro; IPR003016; Lipoyl_BS.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR004099; Pyr_redox_dim.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF00070; Pyr_redox; 1.
DR Pfam; PF02852; Pyr_redox_dim; 1.
DR PRINTS; PR00368; FADPNR.

DR PRINTS; PR00945; HGRDTASE.
DR PRODOM; PR00411; PNDRTASEI.
DR ProDom; PD000139; FAD_pyr_redox; 1.
DR TIGRFAMS; TIGR01350; lipoamide_DH; 1.
DR PROSITE; PS00189; LIPOYL; 1.
DR PROSITE; PS00076; PYRIDINE REDOX; 1.
KW FAD; Flavoprotein; Lipoyl; NAD; Oxidoreductase; Redox-active center.
SQ SEQUENCE 594 AA; 62110 MW; D0E9EC3432B21A44 CRC64;

Query Match 74.5%; Score 35; DB 2; Length 594;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
||| |
Db 233 FLDPHNEV 241

RESULT 8
O83654 PRELIMINARY; PRT; 682 AA.
AC O83654;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein TP0648.
GN OrderedLocustNames=TP0648;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey B.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388 (1998).
DR EMBL; AE001239; AAC65621.1; -.
DR PIR; G71299; G71299.
DR TIGR; TP0648; -.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR; 8.
DR SMART; SM00028; TPR; 9.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Complete proteome; Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 682 AA; 77635 MW; 9C6F0E424EFDCEB CRC64;

Query Match 74.5%; Score 35; DB 2; Length 682;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
:|||||:
Db 437 YLDKALSLI 445

RESULT 9
Q8KJQ3 PRELIMINARY; PRT; 172 AA.
AC Q8KJQ3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical 18.9 kd protein in hpt-pand intergenic region.

GN Name=YadJ;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN SEQUENCE FROM N.A.
RP STRAIN=K-12;
RC MEDLINE=83209630; PubMed=6343085;
RX Stephens P.E., Darlison M.G., Lewis H.M., Guest J.R.;
RA "The pyruvate dehydrogenase complex of Escherichia coli K12.
RT Nucleotide sequence encoding the pyruvate dehydrogenase component.";
RL Eur. J. Biochem. 133:155-162(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=83234434; PubMed=6345153;
RA Stephens P.E., Darlison M.G., Lewis H.M., Guest J.R.;
RT "The pyruvate dehydrogenase complex of Escherichia coli K12.
RT Nucleotide sequence encoding the dihydrolipoamide acetyltransferase
RT component.";
RL Eur. J. Biochem. 133:481-489(1983).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=84004369; PubMed=6352260;
RA Stephens P.E., Lewis H.M., Darlison M.G., Guest J.R.;
RT "Nucleotide sequence of the lipoamide dehydrogenase gene of
RT Escherichia coli K12.";
RL Eur. J. Biochem. 135:519-527(1983).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=87083395; PubMed=3025182;
RA Chye M.L., Pittard J.;
RT "Transcription control of the aroP gene in Escherichia coli K-12:
RT analysis of operator mutants.";
RL J. Bacteriol. 169:386-393(1987).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=85054973; PubMed=6094577;
RA Richard C., Richard F., Martin C., Haziza C., Patte J.C.;
RT "Regulation of expression and nucleotide sequence of the Escherichia
RT coli dapD gene.";
RL J. Biol. Chem. 259:14824-14828(1984).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=86085668; PubMed=3079747;
RA Coulton J.W., Mason P., Cameron D.R., Carmel G., Jean R., Rode H.N.;
RT "Protein fusions of beta-galactosidase to ferrichrome-iron receptor of
RT Escherichia coli K-12.";
RL J. Bacteriol. 165:181-192(1986).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=87014116; PubMed=3020380;
RA Koester W., Braun V.;
RT "Iron hydroxamate transport of Escherichia coli: Nucleotide sequence
RT of the fhuB gene and identification of the protein.";
RL Mol. Gen. Genet. 204:435-442(1986).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=86278132; PubMed=3015933;
RA Breton R., Sanfacon H., Papayannopoulos I., Biemann K., Lapointe J.;
RT "Glutamyl-tRNA synthetase of Escherichia coli. Isolation and primary
RT structure of the gltX gene and homology with other aminoacyl-tRNA
RT synthetases.";
RL J. Biol. Chem. 261:10610-10617(1986).
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=89057448; PubMed=3057437;
RA Lipinska B., Sharma S., Georgopoulos C.;
RT "Sequence analysis and regulation of the htrA gene of Escherichia
RT coli: A sigma-32-independent mechanism of heat-inducible
RT transcription.";
RL Nucleic Acids Res. 16:10053-10067(1988).
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=87109068; PubMed=3027045;
RA Ben-Bassat A., Bauer K., Chang S.Y., Myambo K., Boosman A., Chang S.;
RT "Processing of the initiation methionine from proteins: properties of
RT the Escherichia coli methionine aminopeptidase and its gene
RT structure.";
RL J. Bacteriol. 169:751-757(1987).
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=89155419; PubMed=2537812;
RA Liu J., Parkinson J.S.;
RT "Genetics and sequence analysis of the pcnB locus, an Escherichia coli
RT gene involved in plasmid copy number control.";
RL J. Bacteriol. 171:1254-1261(1989).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=85127060; PubMed=3882429;
RA Broome-Smith J.K., Edelman A., Yousif S., Spratt B.G.;
RT "The nucleotide sequences of the ponA and ponB genes encoding
RT penicillin-binding proteins 1A and 1B of Escherichia coli K12.";
RL Eur. J. Biochem. 147:437-446(1985).
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=82059454; PubMed=6272196;
RA An G., Bendiak D.S., Mamelak L.A., Friesen J.D.;
RT "Organization and nucleotide sequence of a new ribosomal operon in
RT Escherichia coli containing the genes for ribosomal proteins S2 and
RT elongation factors Ts.";
RL Nucleic Acids Res. 9:4163-4172(1981).
RN [14]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=88058963; PubMed=3316212;
RA Tabor C.W., Tabor H.;
RT "The speE operon of Escherichia coli: Formation and processing of
RT a proenzyme form of S-adenosylmethionine decarboxylase.";
RL J. Biol. Chem. 262:16037-16040(1987).
RN [15]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=88152237; PubMed=2450046;
RA Gebhard W., Schreitmuller T., Hochstrasser K.;
RT "Complementary DNA and derived amino acid sequence of the precursor of
RT one of the three protein components of the inter-alpha-trypsin
RT inhibitor complex.";
RL FEBS Lett. 229:63-67(1988).
RN [16]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=88227880; PubMed=3372485;
RA Mellano M.A., Cooksey D.A.;
RT "Nucleotide sequence and organization of copper resistance genes from
RT Pseudomonas syringae pv. tomato.";
RL J. Bacteriol. 170:2879-2883(1988).
RN [17]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=89008347; PubMed=3049588;
RA Sung Y., Fuchs J.A.;
RT "Characterization of the cyn operon in Escherichia coli K12.";
RL J. Biol. Chem. 263:14769-14775(1988).

[18]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K-12;
 RX MEDLINE=90113890; PubMed=2691840;
 RA Lindquist S., Galleni M., Lindberg F., Normark S.;
 RT "Signalling proteins in enterobacterial AmpC beta-lactamase
 regulation.";
 RL Mol. Microbiol. 3:1091-1102(1989).
 [19]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K-12;
 RX MEDLINE=90174991; PubMed=2408019;
 RA Honore N., Cole S.T.;
 RT "Nucleotide sequence of the arcp gene encoding the general aromatic
 amino acid transport protein of Escherichia coli K-12: homology with
 yeast transport proteins.";
 RL Nucleic Acids Res. 18:653-653(1990).
 [20]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K-12;
 RX MEDLINE=89061679; PubMed=2904262;
 RA Andrews S.C., Guest J.R.;
 RT "Nucleotide sequence of the gene encoding the GMP reductase of
 Escherichia coli K12.";
 RL Biochem. J. 255:35-43(1988).
 [21]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K-12;
 RX MEDLINE=90251164; PubMed=2338917;
 RA Surin B.P., Watson J.M., Hamilton W.D., Economou A., Downie J.A.;
 RT "Molecular characterization of the nodulation gene, nodT, from two
 biovars of Rhizobium leguminosarum.";
 RL Mol. Microbiol. 4:245-252(1990).
 [22]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K-12;
 RX MEDLINE=87279948; PubMed=3301821;
 RA Coulton J.W., Mason P., Allatt D.D.;
 RT "fhuC and fhuD genes for iron (III)-ferrichrome transport into
 Escherichia coli K-12.";
 RL J. Bacteriol. 169:3844-3849(1987).
 [23]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K-12;
 RX MEDLINE=89327165; PubMed=2666401;
 RA Xie Q.W., Tabor C.W., Tabor H.;
 RT "Spermidine biosynthesis in Escherichia coli the promoter and the
 termination regions of the speed operon.";
 Query Match 72.3%; Score 34; DB 2; Length 172;
 Best Local Similarity 62.5%; Pred. No. 64;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FLDKAHME 8
 Db 74 FIDKGHLE 81
 RESULT 10
 Q7WS17 PRELIMINARY; PRT; 174 AA.
 AC Q7WS17;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=PHG132, PHG138;
 OS Alcaligenes eutrophus (Ralstonia eutropha).
 OG Plasmid megaplasmid pHG1.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Wautersia.
 OX NCBI_TaxID=510;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=H16; PLASMID=megaplasmid pHG1;
 RX MEDLINE=22830147; PubMed=12948488;
 RA Schwartz E., Henne A., Cramm R., Bitinger T., Friedrich B.,
 RA Gottschalk G.;
 RT "Complete Nucleotide Sequence of pHG1: A Ralstonia eutropha H16
 Megaplasmid Encoding Key Enzymes of H2-based Lithoautotrophy and
 Anaerobiosis.";
 RL J. Mol. Biol. 332:369-383(2003).
 DR EMBL; AY305378; AAP85885.1; -;
 DR EMBL; AY305378; AAP85890.1; -;
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 174 AA; 19544 MW; 50672ACE2C03EE27 CRC64;
 Query Match 72.3%; Score 34; DB 2; Length 174;
 Best Local Similarity 66.7%; Pred. No. 65;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FLDKAHMEV 9
 Db 161 FLDQRHVEV 169
 RESULT 11
 Q9Y0D9 PRELIMINARY; PRT; 220 AA.
 ID Q9Y0D9;
 AC Q9Y0D9;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Transcription factor SoxD1.
 GN Name=SoxD1;
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20025945; PubMed=10556071;
 RA Kenny A.P., Kozlowski D., Oleksyn D.W., Angerer L.M., Angerer R.C.;
 RT "SpSoxBl, a maternally encoded transcription factor asymmetrically
 distributed among early sea urchin blastomeres.";
 RL Development 126:5473-5483(1999).
 DR EMBL; AF157387; AAD40686.1; -;
 DR HSSP; P35710; 1111.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR009071; HMG-box.
 DR InterPro; IPR000910; HMG box; 1.
 DR Pfam; PF00505; HMG box; 1.
 DR SMART; SM00398; HMG; 1.
 DR PROSITE; PS01118; HMG_BOX_2; 1.
 SQ SEQUENCE 220 AA; 25241 MW; 29FE84497C6A866B CRC64;
 Query Match 72.3%; Score 34; DB 2; Length 220;
 Best Local Similarity 85.7%; Pred. No. 85;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LDKAHME 8
 Db 114 LDKAHLE 120
 RESULT 12
 Q9CUS6 PRELIMINARY; PRT; 272 AA.
 ID Q9CUS6;
 AC Q9CUS6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
 library, clone:4833420A15 product:HUNTINGTIN-INTERACTING PROTEIN-1

DE PROTEIN INTERACTOR, full insert sequence. (Fragment).

GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).
RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).
RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA The FANTOM Consortium;

RT "Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).
RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=2049374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).
RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

Kanno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,

Sumi N., Ishii Y., Nakamura S., Hazana M., Nishine T., Harada A.,

Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,

Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).
RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,

Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,

Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,

Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,

Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,

Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,

Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,

Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,

Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,

Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK014731; BAB29522.1; -.

DR MGD; MGI:1921166; Esrrb1.

DR GO; GO:0005794; C:Golgi apparatus; IDA.

FT NON TER 272 272

SQ SEQUENCE 272 AA; 31412 MW; E2A7E9DDF923CF12 CRC64;

Query Match

72.3%; Score 34; DB 2; Length 272;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9
Db 200 FLDKLRNEI 208

RESULT 13

Q6VCY8

ID Q6VCY8 PRELIMINARY; PRT; 275 AA.

AC Q6VCY8;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE MAP1-12.

OS Cowdria ruminantium.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

OC Anaplasmataceae; Ehrlichia.

OX NCBI_TaxID=779;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Weilgevonden;

RA van Heerden H., Collins N.E., Brayton K.A., Rademeyer C.,

RA Allsopp B.A.;

RT "Characterization of a major outer membrane protein multigene family

in Ehrlichia ruminantium.";

RL Gene 330:159-168(2004).

DR EMBL; AY343331; AAR10933.1; -.

DR InterPro; IPR002566; Surface_Ag_msp4.

DR Pfam; PF01617; Surface_Ag_2; 1.

SQ SEQUENCE 275 AA; 31330 MW; 68B5076293032B0C CRC64;

Query Match

72.3%; Score 34; DB 2; Length 275;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9

Db 198 FLDKVFHV 206

RESULT 14

AAR10933

ID AAR10933 PRELIMINARY; PRT; 275 AA.

AC AAR10933;

DT 23-APR-2004 (TREMBLrel. 27, Created)

DT 23-APR-2004 (TREMBLrel. 27, Last sequence update)

DT 23-APR-2004 (TREMBLrel. 27, Last annotation update)

DE MAP1-12.

OS Cowdria ruminantium.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

OC Anaplasmataceae; Ehrlichia.

OX NCBI_TaxID=779;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Weilgevonden;

RA van Heerden H., Collins N.E., Brayton K.A., Rademeyer C.,

RA Allsopp B.A.;

RT "Characterization of a major outer membrane protein multigene family

in Ehrlichia ruminantium.";

RL Gene 330:159-168(2004).

DR EMBL; AY343331; AAR10933.1; -.

SQ SEQUENCE 275 AA; 31330 MW; 68B5076293032B0C CRC64;

Query Match

72.3%; Score 34; DB 2; Length 275;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLDKAHMEV 9

Db 198 FLDKVFHV 206

RESULT 15
ASTE_VIBPA STANDARD; PRT; 342 AA.
AC Q87Q40;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Succinylglutamate desuccinylase (EC 3.1.1.-).
GN Name=astE; OrderedLocusNames=VP1310;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
CC -!- FUNCTION: Transforms N(2)-succinylglutamate into succinate and
glutamate (By similarity).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (Potential).
CC -!- PATHWAY: Arginine catabolism; arginine succinyltransferase
pathway; fifth (last) step.
CC -!- SIMILARITY: Belongs to the aspA/astE family. Succinylglutamate
desuccinylase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005077; BAC59573.1; -.
DR HAMAP; MF_00767; -; 1.
DR InterPro; IPR007036; Aste_AspA.
DR Pfam; PF04952; Aste_AspA; 1.
KW Arginine metabolism; Complete proteome; Hydrolase; Metal-binding;
KW Zinc.
FT ACT_SITE 219 219 Potential.
FT METAL 63 63 Zinc (Potential).
FT METAL 66 66 Zinc (Potential).
FT METAL 155 155 Zinc (Potential).
SQ SEQUENCE 342 AA; 38837 MW; 4FBA79C89D15D5CC CRC64;
Query Match 72.3%; Score 34; DB 1; Length 342;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 FLDKAHME 8
Db 184 FLDSAHIE 191

Search completed: December 14, 2004, 09:19:09
Job time : 10.1135 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31 ; Search time 7.70742 Seconds
(without alignments)
465.433 Million cell updates/sec

Title: US-10-026-001-7

Perfect score: 54

Sequence: 1 FMKKNHEEV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358723299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	54	100.0	10	5	ABB81090	Abb81090 Human cyt
2	54	100.0	35	4	AAM21986	Aam21986 Peptide #
3	54	100.0	35	4	ABB44364	Abb44364 Peptide #
4	54	100.0	35	4	AAM38331	Aam38331 Peptide #
5	54	100.0	35	4	ABB27219	Abb27219 Protein #
6	54	100.0	35	4	AAM78105	Aam78105 Human bon
7	54	100.0	35	4	AAM65428	Aam65428 Human bra
8	54	100.0	35	4	ABG59738	Abg59738 Human liv
9	54	100.0	35	5	ABG47115	Abg47115 Human pep
10	54	100.0	122	5	ABP42263	Abp42263 Human ova
11	54	100.0	135	3	AAB44084	Aab44084 Human can
12	54	100.0	139	8	ABO54622	Abos4622 Human gen
13	54	100.0	209	4	ABG16549	Abgl6549 Novel hum
14	54	100.0	209	4	ABG16692	Abgl6692 Novel hum
15	54	100.0	231	4	ABG09382	Abg09382 Novel hum
16	54	100.0	231	5	ABP42159	Abp42159 Human ova
17	54	100.0	232	5	AAU84288	Aau84288 Human end
18	54	100.0	259	7	ADM05930	Adm05930 Human pro
19	54	100.0	267	4	ABG07825	Abg07825 Novel hum
20	54	100.0	267	8	ADP04216	Adp04216 Human col
21	54	100.0	291	4	ABG27029	Abg27029 Novel hum
22	54	100.0	321	4	ABG00672	Abg00672 Novel hum
23	54	100.0	326	4	ABG15279	Abg15279 Novel hum
24	54	100.0	359	6	AAB38087	Aab38087 Human cyt
25	54	100.0	364	4	ABG27717	Abg27717 Novel hum

26	54	100.0	378	6	ABO14648	Abol14648 Novel hum
27	54	100.0	382	4	ABG15222	Abg15222 Novel hum
28	54	100.0	409	6	ABO14647	Abol14647 Novel hum
29	54	100.0	417	6	AAE33677	Aae33677 Human str
30	54	100.0	423	8	ADF91430	Adf91430 Mouse cyt
31	54	100.0	428	4	ABG08563	Abg08563 Novel hum
32	54	100.0	429	6	AAE38083	Aae38083 Human cyt
33	54	100.0	430	4	AAB90795	Aab90795 Human she
34	54	100.0	430	5	ABB81087	Abb81087 Human cyt
35	54	100.0	430	5	ABP68606	Abp68606 Human pan
36	54	100.0	430	8	ADE76940	Ade76940 Human pro
37	54	100.0	430	8	ADF91429	Adf91429 Human cyt
38	54	100.0	433	4	ABG16694	Abg16694 Novel hum
39	54	100.0	447	4	ABG16550	Abg16550 Novel hum
40	54	100.0	449	4	ABG08564	Abg08564 Novel hum
41	54	100.0	452	4	AAG74328	Aag74328 Human col
42	54	100.0	454	8	ADP04217	Adp04217 Human col
43	54	100.0	456	4	ABG15224	Abg15224 Novel hum
44	51	94.4	312	5	ABB78803	Abb78803 Internedi
45	51	94.4	542	5	ABB78799	Abb78799 Human NOV

ALIGNMENTS

RESULT 1
ABB81090
ID ABB81090 standard; peptide; 10 AA.
XX
AC ABB81090;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human cytokeratin 18 (CK18) compound 3 peptide.
XX
KW Antigen; cytokeratin 18; CK-18; immune response; cytostatic; vaccine;
KW Gene therapy; cancer; human.
XX
OS Homo sapiens.
XX
PN WO200255555-A2.
XX
PD 18-JUL-2002.
XX
PF 21-DEC-2001; 2001WO-US049964.
XX
PR 21-DEC-2000; 2000US-0257820P.
XX
(GENZ) GENZYME CORP.
XX
PI Nicolette CA;
XX
DR WPI; 2002-619103/66.
DR N-PSDB; ABN86613.
XX
PT Novel antigenic cytokeratin 18 compounds and peptides useful for inducing
PT an immune response in a subject and for diagnosing a neoplastic condition
PT or susceptibility to the condition of an animal cell or tissue.
XX
PS Claim 3; Page 63; 73pp; English.
XX
CC The invention relates to novel antigenic cytokeratin 18 (CK-18) compounds
CC (I) and peptides useful for inducing an immune response in a subject. (I)
CC is useful for inducing an immune response in a subject, by delivering (I)
CC as a polynucleotide, in the context of an MHC molecule that presents the
CC compound on the surface of an APC. (I) is useful for aiding in the
CC diagnosis of the neoplastic condition or susceptibility to the condition
CC of an animal cell or tissue; for generating antibodies which are useful
CC for identifying and purifying polypeptides and APCs expressing the
CC polypeptides. (I) serves as markers for the neoplastic phenotype. (I)
CC that is covalently or non-covalently linked to molecules are useful in
CC diagnostic methods, and for detecting or purifying antibodies. It is also
CC useful as components of anti-cancer vaccines and to expand immune

CC effector cells that are specific for cells having differential, i.e.
 CC aberrant, expression of antigenic CK-18. CK-18 proteins or antibodies are
 CC useful for detecting, diagnosing or prognosing and monitoring the
 CC progression, course or stage of CK-18 related cancers, or malignancies.
 CC Host cells comprising one or more immunogenic ligands are useful for
 CC inducing an immune response in a subject, and to expand a population of
 CC immune effector cells such as tumour infiltrating lymphocytes which in
 CC turn are useful in adoptive immunotherapies. Agents that modulate the
 CC binding of CK-18 protein to its ligand are useful for treating disease,
 CC especially cancer. The present sequence represents the human CK18
 CC compound 3 peptide

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 54; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEEV 10
 |||||
 Db 1 FMKKNHEEEV 10

RESULT 2
 AAM21986
 ID AAM21986 standard; protein; 35 AA.

XX AC AAM21986;

XX DT 12-OCT-2001 (first entry)

XX DE Peptide #8420 encoded by probe for measuring cervical gene expression.
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
 XX KW cervical cancer.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US0000670.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX PT gene expression in human cervical epithelial cells.

XX PS Claim 27; SEQ ID NO 26812; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs; see AAI0068-AAI28459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 35 AA;

Query Match 100.0%; Score 54; DB 4; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEEV 10
 |||||
 Db 25 FMKKNHEEEV 34

RESULT 3
 ABB44364

ID ABB44364 standard; peptide; 35 AA.

XX AC ABB44364;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #11870 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US0000669.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX PT gene expression in human fetal liver.

XX PS Claim 27; SEQ ID NO 36999; 639pp + Sequence Listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 35 AA;

Query Match 100.0%; Score 54; DB 4; Length 35;

Best Local Similarity 100.0%; Pred. No. 0.014;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEEV 10
 |||||
 Db 25 FMKKNHEEEV 34

RESULT 4
 AAM38331

ID AAM38331 standard; protein; 35 AA.
XX
AC AAM38331;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #12368 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 38600; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 35 AA;

Query Match 100.0%; Score 54; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10
Db 25 FMKKNHHEEV 34

RESULT 5
ABB27219
ID ABB27219 standard; protein; 35 AA.
XX
AC ABB27219;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #9218 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.

XX 09-AUG-2001.
PD
XX
PF 30-JAN-2001; 2001WO-US0000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 28989; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 35 AA;

Query Match 100.0%; Score 54; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10
Db 25 FMKKNHHEEV 34

RESULT 6
AAM78105
ID AAM78105 standard; protein; 35 AA.
XX
AC AAM78105;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38411.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 XX
 PS Example 4; SEQ ID NO 38411; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 54; DB 4; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FMKKNHHEEV 10
 Db 25 FMKKNHHEEV 34
 RESULT 7
 AAM65428
 ID AAM65428 standard; protein; 35 AA.
 XX
 AC AAM65428;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 37533.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US0000667.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX
 PS Example 4; SEQ ID NO 37533; 650pp + Sequence Listing; English.
 XX

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 54; DB 4; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FMKKNHHEEV 10
 Db 25 FMKKNHHEEV 34
 RESULT 8
 ABG59738
 ID ABG59738 standard; peptide; 35 AA.
 XX
 AC ABG59738;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID NO 38386.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US0000654.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human adult liver.
 PS Claim 27; SEQ ID NO 38386; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SEN) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 35 AA;

Query Match 100.0%; Score 54; DB 4; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEEV 10
 Db 25 FMKKNHEEEV 34

RESULT 9
 ABG47115
 ID ABG47115 standard; peptide; 35 AA.
 XX AC ABG47115;
 XX DT 19-AUG-2002 (first entry)
 XX Human peptide encoded by genome-derived single exon probe SEQ ID 36780.

DE Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

XX OS Homo sapiens.
 XX PN WO200186003-A2.
 XX PD 15-NOV-2001.
 XX PF 30-JAN-2001; 2001WO-US000665.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX Claim 27; SEQ ID NO 36780; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably

labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pot_sequences

XX SQ Sequence 35 AA;
 Query Match 100.0%; Score 54; DB 5; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEEV 10
 Db 25 FMKKNHEEEV 34

RESULT 10
 ABP42263
 ID ABP42263 standard; protein; 122 AA.
 XX AC ABP42263;
 XX DT 22-AUG-2002 (first entry)
 XX Human ovarian antigen HKZAI14, SEQ ID NO:3395.
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX OS Homo sapiens.
 XX PN WO200200677-A1.
 XX PD 03-JAN-2002.
 XX PF 07-JUN-2001; 2001WO-US018569.
 XX PR 07-JUN-2000; 2000US-0209467P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Birse CE, Rosen CA;
 XX WPI; 2002-147878/19.
 XX N-PSDB; ABQ55340.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 11; SEQ ID NO 3395; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides; antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 122 AA;
 Query Match 100.0%; Score 54; DB 5; Length 122;
 Best Local Similarity 100.0%; Pred. No. 0.054;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FMKKNHHEEV 10
 Db 44 FMKKNHHEEV 53
 RESULT 11
 AAB44084
 ID AAB44084 standard; protein; 135 AA.
 AC AAB44084;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1529.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO200053350-A1.

XX 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005882.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587533/55.
 DR N-PSDB; AAC78293.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 XX
 PS Claim 11; Page 2205; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neutropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 135 AA;
 Query Match 100.0%; Score 54; DB 3; Length 135;
 Best Local Similarity 100.0%; Pred. No. 0.06;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FMKKNHHEEV 10
 Db 35 FMKKNHHEEV 44
 RESULT 12
 ABO54622
 ID ABO54622 standard; protein; 139 AA.
 XX
 AC ABO54622;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon protein #856.
 XX
 KW Human; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 OS Homo sapiens.
 XX
 PN US2003194704-A1.
 XX
 PD 16-OCT-2003.
 XX
 PF 03-APR-2002; 2002US-00029386.
 XX

PR 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 45; SEQ ID NO 28256; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 139 AA;

Query Match 100.0%; Score 54; DB 8; Length 139;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10
Db 115 FMKKNHHEEV 124
| | | | | | | | | |

RESULT 13
ABG16549
ID ABG16549 standard; protein; 209 AA.
XX
AC ABG16549;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #16540.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS80736.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 46908; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 209 AA;

Query Match 100.0%; Score 54; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.096;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10
Db 13 FMKKNHHEEV 22
| | | | | | | | | |

RESULT 14
ABG16692
ID ABG16692 standard; protein; 209 AA.
XX
AC ABG16692;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #16683.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS80879.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 47051; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 209 AA;

Query Match 100.0%; Score 54; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.096;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10

Db 13 FMKKNHHEEV 22

RESULT 15

ABG09382

ID ABG09382 standard; protein; 231 AA.

XX AC ABG09382;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #9373.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS73569.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 39741; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 231 AA;

Query Match 100.0%; Score 54; DB 4; Length 231;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10

Db 140 FMKKNHHEEV 149

Search completed: December 14, 2004, 09:13:04
Job time : 8.70742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:31 ; Search time 7.90393 Seconds
(without alignments)
727.960 Million cell updates/sec

Title: US-10-026-001-7
Perfect score: 54
Sequence: 1 FMKNHHEEV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	147	2	Q63278
2	54	100.0	259	2	Q8N7Q0
3	54	100.0	422	1	K1CR_MOUSE
4	54	100.0	429	1	K1CR_HUMAN
5	51	94.4	1110	2	Q91255
6	48	88.9	355	2	Q8N240
7	48	88.9	387	2	Q8AWA6
8	48	88.9	411	1	K1CL_RABIT
9	48	88.9	425	2	Q9PUB6
10	48	88.9	428	2	Q6P0E4
11	48	88.9	428	2	AAH65653
12	48	88.9	429	2	Q9PWB8
13	48	88.9	431	2	Q7SYG0
14	48	88.9	431	2	Q6GLQ7
15	48	88.9	455	2	Q07426
16	48	88.9	456	2	Q6IFW5
17	48	88.9	458	1	K1CM_HUMAN
18	48	88.9	463	2	Q803Y5
19	48	88.9	466	2	Q6DHB6
20	48	88.9	468	1	K1CP_MOUSE
21	48	88.9	473	2	Q9EQD7
22	48	88.9	474	2	Q9EQD6
23	48	88.9	482	2	Q8AWA7
24	48	88.9	483	1	K1CL_MOUSE
25	48	88.9	531	2	Q8AWA8
26	47	87.0	104	2	Q9P1Y3
27	47	87.0	208	2	Q9P1Y4
28	47	87.0	255	1	K1CS_RAT
29	47	87.0	400	1	K1CS_HUMAN
30	47	87.0	400	2	Q96A53
31	47	87.0	401	1	K1CS_POTTR

32 47 87.0 403 1 K1CS_MOUSE
33 47 87.0 403 2 Q6S6J4
34 47 87.0 403 2 AAR36876
35 47 87.0 415 2 Q7LSM9
36 47 87.0 415 2 AAH02539
37 47 87.0 417 2 Q96FV1
38 47 87.0 420 2 Q6NW33
39 47 87.0 420 2 AAH67744
40 47 87.0 438 2 Q9QW74
41 46 85.2 405 2 Q6NYL7
42 46 85.2 405 2 Q6PHE4
43 46 85.2 405 2 AAH65541
44 46 85.2 405 2 AAH56584
45 45 83.3 219 2 Q8N4W2

P19001 mus musculus
Q6s6j4 rattus norv
Aar36876 rattus no
Q7lsm9 homo sapien
Aah02539 homo sapi
Q96fvl homo sapien
Q6nw33 homo sapien
Aah67744 homo sapi
Q90w74 oncorhynchu
Q6nyl7 brachydanio
Q6phe4 brachydanio
Aah65541 brachydian
Aah56584 brachydian
Q8n4w2 homo sapien

ALIGNMENTS

RESULT 1

Q63278 PRELIMINARY; PRT; 147 AA.
AC Q63278;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Keratin 18 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=96081364; PubMed=8541209;
RA Fridmacher V., Le Bert M., Guillou F., Magre S.;
RT "Switch in the expression of the K19/K18 keratin genes as a very early
RT evidence of testicular differentiation in the rat."
RL Mech. Dev. 52:199-207(1995).
DR EMBL; X81448; CAA57204.1; --
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
KW Keratin.
FT NON_TER 1
FT NON_TER 147
SQ SEQUENCE 147 AA; 16960 MW; CE3755F3617FBC08 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 0.096;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKNHHEEV 10
|||
Db 82 FMKNHHEEV 91

RESULT 2

Q8N7Q0 PRELIMINARY; PRT; 259 AA.
AC Q8N7Q0;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein FLJ40504.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=14702039;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura K., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK097823; BAC05178.1; --
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
SQ SEQUENCE 259 AA; 28651 MW; 58EB4F9D6C4F65B3 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FMKQNHHEEV 10
|||||
Db 247 FMKQNHHEEV 256

RESULT 3
K1CR_MOUSE STANDARD; PRT; 422 AA.
AC P05784; Q61766;
DT 01-NOV-1988 (Rel. 09, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Keratin, type I cytoskeletal 18 (Cytokeratin 18) (Cytokeratin endo B)
DE (Keratin D).
GN Name=Krt18; Synonyms=Krt1-18, Kerd;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89196920; PubMed=2467843;
RA Ichinose Y., Morita T., Zhang F., Shimahasonggram S., Tondella M.L.C.,
RA Matsumoto M., Nozaki M., Matsushiro A.;
RT "Nucleotide sequence and structure of the mouse cytoke-
RT gene";
RL Gene 70:85-95(1988).

RN SEQUENCE FROM N.A.
RC TISSUE=Teratocarcinoma;
RA Alonso A., Weber T., Jorcano J.L.;
RT "Cloning and characterization of keratin D, a murine endodermal
RT cytoskeletal protein induced during in vitro differentiation of F9
RT teratocarcinoma cells";
RL Roux's Arch. Dev. Biol. 196:16-21(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86085876; PubMed=2416755;
RA Singer P.A., Trevor K., Oshima R.G.;
RT "Molecular cloning and characterization of the Endo B cytoke-
RT expressed in preimplantation mouse embryos";
RL J. Biol. Chem. 261:538-547(1986).
RN [4]
RP SEQUENCE OF 1-131 FROM N.A.
RX MEDLINE=88255838; PubMed=2454868;
RA Oshima R.G., Trevor K., Shevinsky L.H., Ryder O.A., Cecena G.;
RT "Identification of the gene coding for the Endo B murine cytoke-
RT and its methylated, stable inactive state in mouse nonepithelial
RT cells";
RL Genes Dev. 2:505-516(1988).
CC -1- SUBUNIT: Heterotetramer of two type I and two type II keratins.
CC Keratin 18 associates with keratin 8.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin, I (acidic) and II (neutral to basic) (40-
CC 55 and 56-70 kDa, respectively).
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22832; AAA37552.1; --
CC EMBL; M36376; AAA39373.1; --
CC EMBL; M11686; AAA39390.1; --
CC EMBL; Y00217; CAA68365.1; --
CC PIR; I59463; I59463.
CC HSSP; P08670; IGK7.
CC SWISS-2DPAGE; P05784; MOUSE.
CC MGD; MGI:96692; Krt1-18.
CC InterPro; IPR001664; IF.
CC InterPro; IPR002957; Keratin_I.
CC Pfam; PF00038; Filament; 1.
CC PRINTS; PR01248; TYPE1KERATIN.
CC PROSITE; PS00226; IF; 1.
CC Acetylation; Coiled coil; Glycoprotein; Intermediate filament;
KW Keratin.
FT INIT_MET 0 0 By similarity.
FT MOD_RES 1 1 N-acetylserine (By similarity).
FT DOMAIN 1 70 Head.
FT DOMAIN 71 379 Rod.
FT DOMAIN 380 422 Tail.
FT DOMAIN 71 106 Coiled 1A.
FT DOMAIN 107 124 Linker 1.
FT DOMAIN 125 216 Coiled 1B.
FT DOMAIN 217 240 Linker 12.
FT DOMAIN 241 379 Coiled 2.
FT SITE 263 263 Stutter.
FT SITE 323 323 Stutter.
FT CARBOHYD 30 30 O-linked (GlcNAc) (By similarity).
FT CARBOHYD 31 31 O-linked (GlcNAc) (By similarity).
FT CARBOHYD 49 49 O-linked (GlcNAc) (By similarity).
FT CONFLICT 133 133 L -> F (in Ref. 1).
FT CONFLICT 243 243 D -> N (in Ref. 2).
FT CONFLICT 252 252 A -> G (in Ref. 2).
FT SEQUENCE 422 AA; 47373 MW; 4D5B0E9C7732F2F3 CRC64;
SQ

Query Match 100.0%; Score 54; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKHHEEV 10
DB 204 FMKKHHEEV 213

RESULT 4
K1CR_HUMAN STANDARD; PRT; 429 AA.
AC P05783; Q9BW26;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Keratin, type I cytoskeletal 18 (Cytokeratin 18) (K18) (CK 18).
GN Name-KRT18; Synonyms=CYK18;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87134778; PubMed=2434380;
RA Oshima R.G., Millan J.L., Cecena G.;
RT "Comparison of mouse and human keratin 18: a component of intermediate
RT filaments expressed prior to implantation.";
RL Differentiation 33:61-68(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix, Colon, Placenta, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 6-429 FROM N.A.
RX MEDLINE=87134779; PubMed=2434381;
RA Leube R.E., Bosch F.X., Romano V., Zimbelmann R., Hofler H.,
RA Franke W.W.;
RT "Cytokeratin expression in simple epithelia. III. Detection of mRNAs
RT encoding human cytokeratins nos. 8 and 18 in normal and tumor cells by
RT hybridization with cDNA sequences in vitro and in situ.";
RL Differentiation 33:69-85(1986).
RN [4]
RP SEQUENCE OF 198-429 FROM N.A.
RX MEDLINE=86193258; PubMed=2422083;
RA Romano V., Hatzfeld M., Magin T.M., Zimbelmann R., Franke W.W.,
RA Maier G., Ponstingl H.;
RT "Cytokeratin expression in simple epithelia. I. Identification of mRNA
RT coding for human cytokeratin no. 18 by a cDNA clone.";
RL Differentiation 30:244-253(1986).
RN [5]
RP SEQUENCE OF 1-166 FROM N.A.
RX MEDLINE=88246424; PubMed=2454392;

RA Kulesh D.A., Oshima R.G.;
RT "Cloning of the human keratin 18 gene and its expression in
RT nonepithelial mouse cells.";
RL Mol. Cell. Biol. 8:1540-1550(1988).
RN [6]
RP PARTIAL SEQUENCE.
RC TISSUE=Colon carcinoma;
RX MEDLINE=97295306; PubMed=9150948;
RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins.";
RL Electrophoresis 18:605-613(1997).
RN [7]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=95263524; PubMed=7538124;
RA Ku N.-O., Omary M.B.;
RT "Identification and mutational analysis of the glycosylation sites of
RT human keratin 18.";
RL J. Biol. Chem. 270:11820-11827(1995).
RN [8]
RP VARIANT CRYPTOGENIC CIRRHOSIS LEU-127.
RX MEDLINE=97148766; PubMed=9011570;
RA Ku N.-O., Wright T.L., Terrault N.A., Gish R., Omary M.B.;
RT "Mutation of human keratin 18 in association with cryptogenic
RT cirrhosis.";
RL J. Clin. Invest. 99:19-23(1997).
CC -!- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC Keratin 18 associates with keratin 8.
CC -!- DISEASE: Defects in KRT18 are a cause of cryptogenic cirrhosis
CC [MIM:215600].
CC -!- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin, I (acidic) and II (neutral to basic) (40-
CC 55 and 56-70 kDa, respectively).
CC -!- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF179904; AAA59461.1; -
CC EMBL; BC000180; AAH00180.1; -
CC EMBL; BC000698; AAH00698.1; -
CC EMBL; BC004253; AAH04253.1; -
CC EMBL; BC008636; AAH08636.1; -
CC EMBL; BC020982; AAH20982.1; -
CC EMBL; X12876; CAA31369.1; -
CC EMBL; X12881; CAA31375.1; -
CC EMBL; X12883; CAA31377.1; -
CC PIR; S05481; S05481.
CC HSSP; P08670; IGK7.
CC IntAct; P05783; -
CC GlycoSuiteDB; P05783; -
CC SWISS-2DPAGE; P05783; HUMAN.
CC PMMA-2DPAGE; P05783; -
CC Sienna-2DPAGE; P05783; -
CC Genew; HGNC:6430; KRT18.
CC MIM; 148070; -
CC MIM; 215600; -
CC GO; GO:0005882; C:intermediate filament; TAS.
CC GO; GO:0009653; P:morphogenesis; TAS.
CC InterPro; IPR001664; IF.
CC InterPro; IPR002957; Keratin_I.
CC Pfam; PF00038; Filament; 1.
CC PRINTS; PR01248; TYPE1KERATIN.
CC PROSITE; PS00226; IF; 1.
CC Acetylation; Coiled coil; Direct protein sequencing; Disease mutation;
KW Glycoprotein; Intermediate filament; Keratin.
FT INIT MET 0 0
FT MOD_RES 1 1 N-acetylserine.
FT DOMAIN 1 78 Head.

FT DOMAIN 79 386 Rod.
 FT DOVAIN 387 429 Tail.
 FT DOVAIN 79 114 Coil 1A.
 FT DOVAIN 115 131 Linker 1.
 FT DOVAIN 132 223 Coil 1B.
 FT DOVAIN 224 247 Linker 12.
 FT DOVAIN 248 386 Coil 2.
 FT SITE 270 Stutter.
 FT SITE 330 330 Stutter.
 FT CARBOHYD 29 29 O-linked (GlcNAc).
 FT CARBOHYD 30 30 /FTID=CAR 000175.
 FT CARBOHYD 30 30 O-linked (GlcNAc).
 FT CARBOHYD 48 48 /FTID=CAR 000193.
 FT CARBOHYD 48 48 O-linked (GlcNAc).
 FT VARIANT 127 127 /FTID=CAR 000194.
 FT H -> L (in cryptogenic cirrhosis;
 FT interfers with the ability to form normal
 FT filaments).
 FT /FTID=VAR 003852.
 FT CONFLICT 167 167 Y -> H (in Ref. 2; AAH00698).
 FT CONFLICT 201 201 E -> Q (in Ref. 4).
 FT CONFLICT 245 245 A -> S (in Ref. 4).
 FT CONFLICT 308 308 D -> R (in Ref. 4).
 FT CONFLICT 311 311 S -> R (in Ref. 4).
 FT SEQUENCE 429 AA; 47926 MW; 7172E9D6A39933F8 CRC64;
 SQ

Query Match 100.0%; Score 54; DB 1; Length 429;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
 DB 211 FMKKNHEEV 220

RESULT 5
 Q91255 PRELIMINARY; PRT; 1110 AA.
 AC Q91255
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE NF-180.
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Central nervous system;
 RX MEDLINE=95287814; PubMed=7770000;
 RA Jacobs A.J., Kamholz J., Selzer M.E.;
 RT "The single lamprey neurofilament subunit (NF-180) lacks
 RT multiphosphorylation repeats and is expressed selectively in
 RT projection neurons.";
 RL Brain Res. Mol. Brain Res. 29:43-52(1995).
 CC -1- SIMILARITY: Belongs to the intermediate filament family.
 DR ENBL; U19361; AAA80106.1; -.
 DR PIR; I51116; I51116.
 DR HSP; P08670; IGK7.
 DR GO; GO:0005882; C:intermediate filament; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR006821; Filament_head.
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF00038; Filament; 1.
 DR Pfam; PF04732; Filament_head; 1.
 DR PRINTS; PR00194; TROPOMYOSIN.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament.
 SQ SEQUENCE 1110 AA; 123817 MW; 6558DA73DAF6974C CRC64;

Query Match 94.4%; Score 51; DB 2; Length 1110;

Best Local Similarity 90.0%; Pred. No. 2.7;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
 DB 235 FLKKNHEEV 244

RESULT 6
 Q8N240 PRELIMINARY; PRT; 355 AA.
 AC Q8N240
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein FLJ34957.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Tongue;
 RX PubMed=14702039;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshihawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Serba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR ENBL; AK092276; BAC03847.1; -.
 DR HSP; P08670; IGK7.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR002957; Keratin_I.
 DR Pfam; PF00038; Filament; 1.
 DR PRINTS; PR01248; TYPE1KERATIN.
 SQ SEQUENCE 355 AA; 38579 MW; C177447C2A7AC3A3 CRC64;

Query Match 88.9%; Score 48; DB 2; Length 355;
 Best Local Similarity 80.0%; Pred. No. 2.9;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
 DB 226 YMKKNHEEV 235


```

RESULT 7
Q8AWA6 ID Q8AWA6 PRELIMINARY; PRT; 387 AA.
AC Q8AWA6;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Type I keratin 18.
OS Lampetra fluviatilis (River lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7748;
RN [1]
RP SEQUENCE FROM N.A.
RA Schultess J., Schaffeld M., Markl J.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; AJ308118; CAC87098.1; -.
DR HSSP; P08670; 1GK7.
DR GO; GO:0005982; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Keratin.
SQ SEQUENCE 387 AA; 42613 MW; 636248E3C5EC6CC0 CRC64;

Query Match 88.9%; Score 48; DB 2; Length 387;
Best Local Similarity 80.0%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
Db 196 FLKKNHEEL 205

RESULT 8
K1CL RABIT ID K1CL RABIT STANDARD; PRT; 411 AA.
AC Q28706;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Keratin, type I cytoskeletal 12 (Cytokeratin 12) (Fragment).
GN Name=KRT12;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cornea;
RX MEDLINE=94192891; PubMed=7511548;
RA Wu R.-L., Zhu G., Galvin S., Xu C., Haseba T., Chaloin-Dufau C.,
RA Dhoulailly D., Wei Z.-G., Lavker R.M., Kao W.-Y., Sun T.-T.;
RT "Lineage-specific and differentiation-dependent expression of K12
RT keratin in rabbit corneal/limbal epithelial cells: cDNA cloning and
RT northern blot analysis."
RL Differentiation 55:137-144 (1994).
CC -1- FUNCTION: May play a unique role in maintaining the normal corneal
CC epithelial function (By similarity).
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC Keratin 3 associates with keratin 12.
CC -1- TISSUE SPECIFICITY: Cornea specific. Associated mainly with all
CC layers of the central corneal epithelium and also found in the
CC suprabasal limbal epithelium.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin, I (acidic) and II (neutral to basic) (40-
CC 55 and 56-70 kDa, respectively).
CC -1- SIMILARITY: Belongs to the intermediate filament family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X77665; CAA54741.1; -.
DR PIR; S45318; S45318.
DR HSSP; P08670; 1GK7.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PROSITE; PS00226; IF; 1.
KW Coiled coil; Intermediate filament; Keratin.
FT NON_TER 1
FT DOMAIN <1 42 Head.
FT DOMAIN 43 354 Rod.
FT DOMAIN 355 411 Tail.
FT DOMAIN 43 78 Coil 1A.
FT DOMAIN 83 101 Linker 1.
FT DOMAIN 102 193 Coil 1B.
FT DOMAIN 194 216 Linker 12.
FT DOMAIN 217 354 Coil 2.
SQ SEQUENCE 411 AA; 45726 MW; 0B22E9A4D9DC250 CRC64;

Query Match 88.9%; Score 48; DB 1; Length 411;
Best Local Similarity 80.0%; Pred. No. 3.4;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
Db 181 YMKKNHEEL 190

RESULT 9
Q9PUB6 ID Q9PUB6 PRELIMINARY; PRT; 425 AA.
AC Q9PUB6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Type I cytokekeratin.
GN Name=cki; Synonyms=CKI;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Chua K.L., Lim T.M.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197880; AAF05846.1; -.
DR HSSP; P08670; 1GK7.
DR ZFIN; ZDB-GENE-991110-22; cki.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
SQ SEQUENCE 425 AA; 46721 MW; 2A04A85AA9A8407B CRC64;

Query Match 88.9%; Score 48; DB 2; Length 425;
Best Local Similarity 80.0%; Pred. No. 3.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
Db 211 FLKKNHEEL 220

```



```

RESULT 10
Q6POE4
ID Q6POE4 PRELIMINARY; PRT; 428 AA.
AC Q6POE4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type I cytochrome, enveloping layer.
GN Name=cytl;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065653; AAH65653.1; -.
SQ SEQUENCE 428 AA; 46553 MW; C581D007F6EBA0EB CRC64;

Query Match 88.9%; Score 48; DB 2; Length 428;
Best Local Similarity 80.0%; Pred. No. 3.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
DB 214 FLKKNHEEL 223

RESULT 11
AAH65653
ID AAH65653 PRELIMINARY; PRT; 428 AA.
AC AAH65653;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Type I cytochrome, enveloping layer.
GN Cyt1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.

Query Match 88.9%; Score 48; DB 2; Length 429;
Best Local Similarity 80.0%; Pred. No. 3.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
DB 214 FLKKNHEEL 223

RESULT 12
Q9PWD8
ID Q9PWD8 PRELIMINARY; PRT; 429 AA.
AC Q9PWD8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Type I cytochrome.
GN Name=cytl;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Sagerstrom C.G., Gammill L.S., Veale R.V., Sive H.L.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF084461; AAD52042.1; -.
DR HSSP; P08670; 1GK7.
DR ZFIN; ZDB-GENE-991008-6; cyt1.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
SQ SEQUENCE 429 AA; 46684 MW; 816CC018A87922A9 CRC64;

Query Match 88.9%; Score 48; DB 2; Length 429;
Best Local Similarity 80.0%; Pred. No. 3.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
DB 214 FLKKNHEEL 223

```

QY 1 FMKKNHEEV 10
:|||||:
Db 215 FLKKNHEEL 224

RESULT 13

Q7SYGO PRELIMINARY; PRT; 431 AA.
AC Q7SYGO;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Keratin type IE.
GN Name=krt14;
OS Acipenser baerii (Siberian sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=27689;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mixture of gill;
RA Haberkamp M., Schaffeld M., Markl J.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; AJ493259; CAD38122.1; -.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Keratin.
SQ SEQUENCE 431 AA; 46759 MW; 6C164C26393412B9 CRC64;

Query Match 88.9%; Score 48; DB 2; Length 431;
Best Local Similarity 80.0%; Pred. No. 3.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
:|||||:
Db 226 FLKKNHEEL 235

RESULT 14

Q6GLQ7 PRELIMINARY; PRT; 431 AA.
AC Q6GLQ7;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; BC074402; AAH74402.1; -.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin I.
DR InterPro; IPR009053; Prefoldin.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPEIKERATIN.
DR PROSITE; PS00226; IF; UNKNOWN 1.
KW Hypothetical protein; Intermediate filament.
SQ SEQUENCE 431 AA; 47723 MW; 0032E596D1B243C8 CRC64;

Query Match 88.9%; Score 48; DB 2; Length 431;
Best Local Similarity 80.0%; Pred. No. 3.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
:|||||:
Db 218 YMKKNHEEM 227

RESULT 15

Q07426 PRELIMINARY; PRT; 455 AA.
AC Q07426;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Keratin.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Optic nerve;
RX MEDLINE=94259853; PubMed=7515399;
RA Druger R.K., Glasgow E., Fuchs C., Levine E.M., Matthews J.P.,
RA Park C.Y., Schechter N.;
RT "Complex expression of keratins in goldfish optic nerve.";
RL J. Comp. Neurol. 340:269-280 (1994).
DR EMBL; I09743; AAC38008.1; -.
DR HSSP; P08670; 1GK7.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin I.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PR01248; TYPEIKERATIN.
SQ SEQUENCE 455 AA; 49755 MW; 8C1378F2F3848EA4 CRC64;

Query Match 88.9%; Score 48; DB 2; Length 455;

Best Local Similarity 80.0%; Pred. No. 3.7;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PMKNHEEV 10
|:|||||:
Db 238 FLKNHEEL 247

Search completed: December 14, 2004, 09:19:11
Job time: 9.90393 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:43:49 ; Search time 1.52838 Seconds
(without alignments)
629.533 Million cell updates/sec

Title: US-10-026-001-7
Perfect score: 54
Sequence: 1 FMKKNHEEV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	423	2 I59463	keratin, type I, c
2	54	100.0	430	2 S05481	keratin 18, type I
3	51	94.4	1110	2 I51116	NF-180 - sea lamp
4	48	88.9	411	2 S45318	keratin 12 - rabbi
5	48	88.9	420	2 A37343	keratin 13, type I
6	48	88.9	458	1 KRHU3	keratin 13, type I
7	48	88.9	483	2 A55033	keratin 12 - mouse
8	47	87.0	400	1 KRHU9	keratin 19, type I
9	47	87.0	401	2 S57657	keratin 19 - potot
10	47	87.0	403	2 JQ0028	cytokekeratin 19 - m
11	45	83.3	359	2 B26135	keratin, 52K type
12	45	83.3	392	2 A60777	keratin 2, type I,
13	45	83.3	399	2 A25470	cytokekeratin 19 - b
14	45	83.3	419	2 A25438	keratin, type I cy
15	45	83.3	429	2 A25145	keratin, 47K type
16	45	83.3	432	2 S30433	keratin 17, type I
17	45	83.3	433	2 S01631	keratin, type I, e
18	45	83.3	437	2 A55682	keratin 13, type I
19	45	83.3	452	2 I49595	cytokekeratin 15 - m
20	45	83.3	456	1 KRHU5	keratin 15, type I
21	45	83.3	467	2 I50476	keratin type I - g
22	45	83.3	472	1 KRHUE	keratin 14, type I
23	45	83.3	526	1 KRBOVI	keratin, 54K type
24	45	83.3	569	1 KRME1	keratin, 59K type
25	45	83.3	593	1 KRHU0	keratin 10, type I
26	45	83.3	849	2 S00030	neurofilament trip
27	45	83.3	913	2 T52485	neurofilament prot
28	43	79.6	532	1 QRPQM	neurofilament trip
29	43	79.6	784	2 PNO009	neurofilament trip

ALIGNMENTS

RESULT 1

I59463
keratin, type I, cytoskeletal - mouse
N;Alternate names: endo B cytokekeratin; keratin D
C;Species: Mus musculus (house mouse)
C;Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 09-Jul-2004
C;Accession: I59463; A25621; A28428; JT0406
R;Alonso, A.; Weber, T.; Jorcano, J.L.
Roux's Arch. Dev. Biol. 196, 16-21, 1987
A;Title: Cloning and characterization of keratin D, a murine endodermal cytoskeletal pro
A;Reference number: I59463
A;Accession: I59463
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-423 <RES>
A;Cross-references: UNIPROT:P05784; GB:M36376; NID:G198587; PIDN:AAA39373.1; PID:G293682
R;Singer, P.A.; Trevor, K.; Oshima, R.G.
J. Biol. Chem. 261, 538-547, 1986
A;Title: Molecular cloning and characterization of the endo B cytokekeratin expressed in p
A;Reference number: A25621; MUID:86085876; PMID:2416755
A;Accession: A25621
A;Molecule type: mRNA
A;Residues: 1-243, 'D', 245-252, 'A', 254-423 <SIN>
A;Cross-references: GB:M11686; NID:G198620; PIDN:AAA39390.1; PID:G293685
R;Oshima, R.G.; Trevor, K.; Shevinsky, L.H.; Ryder, O.A.; Cecena, G.
Genes Dev. 2, 505-516, 1988
A;Title: Identification of the gene coding for the endo B murine cytokekeratin and its met
A;Reference number: A28428; MUID:88255838; PMID:2454868
A;Accession: A28428
A;Molecule type: DNA
A;Residues: 1-132 <OSH>
A;Cross-references: GB:Y00217; NID:G50842; PIDN:CAA68365.1; PID:G50843
R;Ichinose, Y.; Morita, T.; Zhang, F.; Srimahonggram, S.; Tondella, M.L.C.; Matsumoto,
Gene 70, 85-95, 1988
A;Title: Nucleotide sequence and structure of the mouse cytokekeratin endoB gene.
A;Reference number: JT0406; MUID:89198920; PMID:2467843
A;Accession: JT0406
A;Molecule type: DNA
A;Residues: 1-133, 'F', 135-243, 'D', 245-252, 'A', 254-423 <ICH>
A;Cross-references: GB:M22832; NID:G340757; PIDN:AAA37552.1; PID:G532610
C;Genetics:
A;Gene: endoB; KERG
A;Introns: 132/3; 160/2; 212/3; 267/3; 309/3; 384/2
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; intermediate filament

Query Match 100.0%; Score 54; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
|||||||

Db 205 FMKKNHEEV 214

RESULT 2

S05481
keratin 18, type I, cytoskeletal - human
N;Alternate names: cytokeratin 18
C;Species: Homo sapiens (man)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S05481; S05482; S06889
R;Oshima, R.G.; Millan, J.L.; Cecena, G.
Differentiation 33, 61-68, 1986
A;Title: Comparison of mouse and human keratin 18: a component of intermediate filaments
A;Reference number: S05481; MUID:87134778; PMID:2434380
A;Accession: S05481
A;Molecule type: mRNA
A;Residues: 1-430 <OSH>
A;Cross-references: UNIPROT:P05783; EMBL:X12881; NID:G34036; PIDN:CAA31375.1; PID:G34037
R;Romano, V.; Hatzfeld, M.; Magin, T.M.; Zimbelmann, R.; Franke, W.W.; Maier, G.; Ponstl
Differentiation 30, 244-253, 1986
A;Title: Cytokeratin expression in simple epithelia. I. Identification of mRNA coding for
A;Reference number: S05482; MUID:86193258; PMID:2422083
A;Accession: S05482
A;Molecule type: mRNA
A;Residues: 199-201, 'Q', 203-245, 'S', 247-308, 'R', 310-311, 'R', 313-430 <ROM>
A;Cross-references: EMBL:X12876; NID:G34034; PIDN:CAA31369.1; PID:G34035
A;Note: part of this sequence was confirmed by protein sequencing
R;Leube, R.E.; Bosch, F.X.; Romano, V.; Zimbelmann, R.; Hoefler, H.; Franke, W.W.
Differentiation 33, 69-85, 1986
A;Title: Cytokeratin expression in simple epithelia.
A;Reference number: S06888; MUID:87134779; PMID:2434381
A;Accession: S06889
A;Molecule type: mRNA
A;Residues: 7-430 <LEU>
A;Cross-references: EMBL:X12883; NID:G30310; PIDN:CAA31377.1; PID:G30311
C;Genetics:
A;Gene: GDB:KRT18
A;Cross-references: GDB:120127; OMIM:148070
A;Map position: 17p12-17p11
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; intermediate filament
F;2-430/Product: keratin 18, type I, cytoskeletal #status predicted <MAT>

Query Match 100.0%; Score 54; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10

Db 212 FMKKNHEEV 221

RESULT 3

IS1116
NF-180 - sea lamprey
C;Species: Petromyzon marinus (sea lamprey)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: IS1116
R;Jacobs, A.J.; Kamholz, J.; Selzer, M.E.
Brain Res. Mol. Brain Res. 29, 43-52, 1995
A;Title: The single lamprey neurofilament subunit (NF-180) lacks multiphosphorylation re
A;Reference number: IS1116; MUID:95287814; PMID:7770000
A;Accession: IS1116
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1110 <JAC>
A;Cross-references: UNIPROT:Q91255; EMBL:U19361; NID:G632548; PIDN:AAA80105.1; PID:G6325
C;Superfamily: neurofilament triplet H protein

Query Match 94.4%; Score 51; DB 2; Length 1110;
Best Local Similarity 90.0%; Pred. No. 0.35;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10

Db 235 FMKKNHEEV 244

RESULT 4

S45318
keratin 12 - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 06-Jan-1995 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: S45318
R;Wu, R.L.; Zhu, G.; Galvin, S.; Xu, C.; Haseba, T.; Chaloin-Dufau, C.; Dhouailly, D.; We
Differentiation 55, 137-144, 1994
A;Title: Lineage-specific and differentiation-dependent expression of K12 keratin in rab
A;Reference number: S45318; MUID:94192891; PMID:7511548
A;Accession: S45318
A;Molecule type: mRNA
A;Residues: 1-411 <WUR>
A;Cross-references: UNIPROT:Q28706; EMBL:X77665; NID:G495260; PIDN:CAA54741.1; PID:G49526
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; intermediate filament

Query Match 88.9%; Score 48; DB 2; Length 411;
Best Local Similarity 80.0%; Pred. No. 0.46;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10

Db 181 YMKKNHEEL 190

RESULT 5

A37343
keratin 13, type I, cytoskeletal, short splice form - human
N;Alternate names: cytokeratin 13
C;Species: Homo sapiens (man)
C;Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 09-Jul-2004
C;Accession: A37343
R;Kuruc, N.; Leube, R.E.; Moll, I.; Bader, B.L.; Franke, W.W.
Differentiation 42, 111-123, 1989
A;Title: Synthesis of cytokeratin 13, a component characteristic of internal stratified e
A;Reference number: A37343; MUID:90228645; PMID:2483837
A;Accession: A37343
A;Molecule type: mRNA
A;Residues: 1-420 <KUR>
A;Cross-references: UNIPROT:P13646; GB:X52426; NID:G30376; PIDN:CAA36673.1; PID:G30377
C;Genetics:
A;Gene: GDB:KRT13
A;Cross-references: GDB:120740; OMIM:148065
A;Map position: 17q21-17q22
C;Superfamily: cytoskeletal keratin
C;Keywords: alternative splicing; coiled coil; intermediate filament
F;1-96/Domain: head <HEA>
F;97-412/Domain: helical rod #status predicted <ROD>
F;413-420/Domain: tail <TAI>

Query Match 88.9%; Score 48; DB 2; Length 420;
Best Local Similarity 80.0%; Pred. No. 0.47;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10

Db 238 YMKKNHEEM 247

RESULT 6

KRHU3
keratin 13, type I, cytoskeletal, long splice form - human
N;Alternate names: cytokeratin 13
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S06088; A3216; B37343; A33403
R;Mischke, D.; Wächter, E.; Hochstrasser, K.; Wild, A.G.; Schulz, P.

Nucleic Acids Res. 17, 7984, 1989
 A;Title: The N-, but not the C-terminal domains of human keratins 13 and 15 are closely
 A;Reference number: S06086; MUID:90016882; PMID:2477803
 A;Accession: S06088
 A;Molecule type: mRNA
 A;Residues: 1-458 <MIS2>
 A;Cross-references: UNIPROT:P13646; EMBL:X14640; NID:G34032; PIDN:CAA32786.1; PID:G34033
 A;Accession: A33216
 A;Molecule type: protein
 A;Residues: 291-299 <MIS2>
 R;Kuruc, N.; Leube, R.E.; Moll, I.; Bader, B.L.; Franke, W.W.
 Differentiation 42, 111-123, 1989
 A;Title: Synthesis of cytokeratin 13, a component characteristic of internal stratified
 A;Reference number: A37343; MUID:90228645; PMID:2483837
 A;Accession: B37343
 A;Molecule type: mRNA
 A;Residues: 1-57, 'G', 59-458 <KUR>
 A;Cross-references: GB:X52426
 R;Schulz, P.; Wächter, E.; Hochstrasser, K.; Wild, A.G.; Mischke, D.
 Biochem. Biophys. Res. Commun. 162, 1522-1527, 1989
 A;Title: Sequence of a human keratin 13 specific cDNA encompassing coil 1B through the 3
 A;Reference number: A33403; MUID:89350978; PMID:2475110
 A;Accession: A33403
 A;Molecule type: mRNA
 A;Residues: 158-458 <SCH>
 C;Genetics:
 A;Gene: GDB:KRT13
 A;Cross-references: GDB:120740; OMIM:148065
 A;Map position: 17q21-17q22
 C;Superfamily: cytoskeletal keratin
 C;Keywords: alternative splicing; coiled coil; intermediate filament
 F;1-96/Domain: head <HEA>
 F;97-412/Domain: helical rod #status predicted <ROD>
 F;413-458/Domain: tail <TAI>

Query Match 88.9%; Score 48; DB 1; Length 458;
 Best Local Similarity 80.0%; Pred. No. 0.51;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
 :|||||:
 Db 238 YMKNHEEM 247

RESULT 7
 A55033
 keratin 12 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C;Accession: A55033
 R;Liu, C.Y.; Zhu, G.; Converse, R.; Kao, C.W.C.; Nakamura, H.; Tseng, S.C.G.; Mui, M.M.;
 J. Biol. Chem. 269, 24627-24636, 1994
 A;Title: Characterization and chromosomal localization of the cornea-specific murine ker
 A;Reference number: A55033; MUID:95014223; PMID:7523376
 A;Accession: A55033
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-483 <LIU>
 A;Cross-references: UNIPROT:Q64291; GB:U08095; NID:G565659; PIDN:AA52359.1; PID:G565660
 A;Note: authors translated the codon ATC for residue 225 as Thr, and GCG for residue 388
 C;Superfamily: cytoskeletal keratin

Query Match 88.9%; Score 48; DB 2; Length 483;
 Best Local Similarity 80.0%; Pred. No. 0.54;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
 :|||||:
 Db 255 YMKNHEEL 264

RESULT 8
 KRHU9

keratin 19, type I, cytoskeletal - human
 N;Alternate names: cytokeratin 19
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1991 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
 C;Accession: A31370; A61556; A60779; S00658; S60152
 R;Eckert, R.L.
 Proc. Natl. Acad. Sci. U.S.A. 85, 1114-1118, 1988
 A;Title: Sequence of the human 40-kDa keratin reveals an unusual structure with very high
 A;Reference number: A31370; MUID:88124986; PMID:2448790
 A;Accession: A31370
 A;Molecule type: mRNA
 A;Residues: 1-400 <ECK>
 A;Cross-references: UNIPROT:Q96A53; GB:J03607; NID:G184658; PIDN:AAA36044.1; PID:G386803
 R;Bader, B.L.; Jahn, L.; Franke, W.W.
 Eur. J. Cell Biol. 47, 300-319, 1988
 A;Title: Low level expression of cytokeratins 8, 18 and 19 in vascular smooth muscle cel
 containing the cytokeratin 19 gene.
 A;Reference number: A61556; MUID:89210901; PMID:2468493
 A;Accession: A61556
 A;Molecule type: mRNA
 A;Residues: 1-400 <BAD>
 R;Stasiak, P.C.; Purkis, P.E.; Leigh, I.M.; Lane, E.B.
 J. Invest. Dermatol. 92, 707-716, 1989
 A;Title: Keratin 19: predicted amino acid sequence and broad tissue distribution suggest
 A;Reference number: A60779; MUID:89235250; PMID:2469734
 A;Accession: A60779
 A;Molecule type: mRNA
 A;Residues: 1-349, 'A', 351-400 <STA>
 A;Cross-references: EMBL:Y00503; NID:G34038; PIDN:CAA68556.1; PID:G34039
 R;Stasiak, P.C.; Lane, E.B.
 Nucleic Acids Res. 15, 10058, 1987
 A;Title: Sequence of cDNA coding for human keratin 19.
 A;Reference number: S00658; MUID:88096504; PMID:2447559
 A;Accession: S00658
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-75, 'H', 78-341, 'Y', 343-349, 'A', 351-400 <ST2>
 A;Cross-references: EMBL:Y00503
 A;Note: this sequence has been revised
 R;Petersen, G.; Song, D.; Huegle-Doerr, B.; Oldenburg, I.; Bautz, E.K.F.
 Mol. Gen. Genet. 249, 425-431, 1995
 A;Title: Mapping of linear epitopes recognized by monoclonal antibodies with gene-fragment
 A;Reference number: S60151; MUID:96133682; PMID:8552047
 A;Accession: S60152
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 331-349, 'A', 351-380 <PET>
 C;Comment: Keratin 19 is the smallest human keratin, as the non-alpha-helical tail preser
 C;Genetics:
 A;Gene: GDB:KRT19
 A;Cross-references: GDB:120131; OMIM:148020
 A;Map position: 17q21-17q23
 C;Superfamily: cytoskeletal keratin
 C;Keywords: coiled coil; intermediate filament
 F;1-73/Domain: head <HEA>
 F;74-386/Domain: helical rod #status predicted <ROD>
 F;387-399/Domain: tail <TAI>

Query Match 87.0%; Score 47; DB 1; Length 400;
 Best Local Similarity 70.0%; Pred. No. 0.67;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
 :|||||:
 Db 213 YLKNHEEI 222

RESULT 9
 S57657

keratin 19 - potoroo
 C;Species: Potorous tridactylus (potoroo, long-nosed rat kangaroo)
 C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
 C;Accession: S66257; S57657

R;Boettger, V.; Stasiak, P.C.; Harrison, D.L.; Mellerick, D.M.; Lane, E.B.
Eur. J. Biochem. 231, 475-485, 1995
A;Title: Epitope mapping of monoclonal antibodies to keratin 19 using keratin fragments,
A;Reference number: S66257; MUID:95361872; PMID:7543411
A;Accession: S66257
A;Molecule type: mRNA
A;Residues: 1-401 <BOE>
A;Cross-references: UNIPROT:P51856; EMBL:X82579; NID:9886903; PIDN:CAA57915.1; PID:98869
C;Genetics:
A;Gene: KRT19
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; intermediate filament
F;1-73/Domain: head <HEA>
F;74-386/Domain: helical rod #status predicted <ROD>

Query Match 87.0%; Score 47; DB 2; Length 401;
Best Local Similarity 70.0%; Pred. No. 0.68;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
:|||||:
Db 213 YLKNHEEI 222

RESULT 10

JQ0028
cyto keratin 19 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: JQ0028; JQ0378; JN0101
R;Ichinose, Y.; Hashido, K.; Miyamoto, H.; Nagata, T.; Nozaki, M.; Morita, T.; Matsushita
Gene 80, 315-323, 1989
A;Title: Molecular cloning and characterization of cDNA encoding mouse cyto keratin No. 1
A;Reference number: JQ0028; MUID:90060780; PMID:2479589
A;Accession: JQ0028
A;Molecule type: mRNA
A;Residues: 1-403 <ICH>
A;Cross-references: UNIPROT:P19001; GB:M28698; NID:G623167; PIDN:AAA60432.1; PID:G623168
R;Lusier, M.; Ouellet, T.; Lampron, C.; Lapointe, L.; Royal, A.
Gene 85, 435-444, 1989
A;Title: Mouse keratin 19: complete amino acid sequence and gene expression during devel
A;Reference number: JQ0378; MUID:90185218; PMID:2483396
A;Accession: JQ0378
A;Molecule type: mRNA
A;Residues: 1-403 <LUS>
R;Lusier, M.; Filion, M.; Compton, J.G.; Nadeau, J.H.; Lapointe, L.; Royal, A.
Gene 95, 203-213, 1990
A;Title: The mouse keratin 19-encoding gene: sequence, structure and chromosomal assignm
A;Reference number: JN0101; MUID:91065533; PMID:1701153
A;Accession: JN0101
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <LUS2>
A;Cross-references: GB:M36120; NID:G198583; PIDN:AAA39371.1; PID:G387393
C;Comment: Cyto keratin polypeptides are major components of intermediate filaments which
C;Genetics:
A;Gene: K19
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil
F;1-75/Domain: head <HED>
F;76-118/Domain: coil 1a, alpha-helical rod <C1A>
F;133-228/Domain: coil 1b, alpha-helical rod <C1B>
F;249-390/Domain: coil 2, alpha-helical rod <HC2>

Query Match 87.0%; Score 47; DB 2; Length 403;
Best Local Similarity 70.0%; Pred. No. 0.68;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
:|||||:
Db 216 YLKNHEEI 225

RESULT 11

B26135
keratin, 52K type I epidermal (clone pKSCC 52) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jul-2004
C;Accession: B26135
R;Knapp, B.; Rentrop, M.; Schweitzer, J.; Winter, H.
J. Biol. Chem. 262, 938-945, 1987
A;Title: Three cDNA sequences of mouse type I keratins. Cellular localization of the mRNA
A;Reference number: A92654; MUID:87109202; PMID:2433272
A;Accession: B26135
A;Molecule type: mRNA
A;Residues: 1-359 <KNA>
A;Cross-references: UNIPROT:Q61781; GB:M13806; NID:G198628; PIDN:AAA39392.1; PID:G387398
C;Superfamily: cytoskeletal keratin

Query Match 83.3%; Score 45; DB 2; Length 359;
Best Local Similarity 70.0%; Pred. No. 1.4;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
:|||||:
Db 129 YLKNHEEM 138

RESULT 12

A60777
keratin 2, type I, hair - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 09-Jul-2004
C;Accession: A60777
R;Bertolino, A.P.; Checkla, D.M.; Heitner, S.; Freedberg, I.M.; Yu, D.
J. Invest. Dermatol. 94, 297-303, 1990
A;Title: Differential expression of type I hair keratins.
A;Reference number: A60777; MUID:90171610; PMID:1689759
A;Accession: A60777
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-392 <BER>
A;Cross-references: UNIPROT:Q9D646
C;Superfamily: cytoskeletal keratin
C;Keywords: hair

Query Match 83.3%; Score 45; DB 2; Length 392;
Best Local Similarity 88.9%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MKKNHEEV 10
:|||||:
Db 190 LKKNHEEV 198

RESULT 13

A25470
cyto keratin 19 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C;Accession: A25470
R;Bader, B.L.; Magin, T.M.; Hatzfeld, M.; Franke, W.W.
EMBO J. 5, 1865-1875, 1986
A;Title: Amino acid sequence and gene organization of cyto keratin no. 19, an exceptional
A;Reference number: A25470; MUID:87004553; PMID:2428612
A;Accession: A25470
A;Molecule type: DNA
A;Residues: 1-399 <BAD>
A;Cross-references: UNIPROT:P08728; GB:X04152; NID:G469; PIDN:CAA27770.1; PID:G1197196
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 83.3%; Score 45; DB 2; Length 399;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FMKKNHEEV 10
:|||||:
Db 212 YLKNHEEM 221

RESULT 14

A25438
keratin, type I cytoskeletal protein, B2 - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C;Accession: A25438
R;Miyatani, S.; Winkles, J.A.; Sargent, T.D.; Dawid, I.B.
J. Cell Biol. 103, 1957-1965, 1986
A;Title: Stage-specific keratins in Xenopus laevis embryos and tadpoles: the XK81 gene f
A;Reference number: A25438; MUID:87057649; PMID:2430981
A;Accession: A25438
A;Molecule type: mRNA
A;Residues: 1-419 <MIY>
A;Cross-references: UNIPROT:P05781; GB:M18155; GB:X05865; NID:g214558; PID:g214560
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 83.3%; Score 45; DB 2; Length 419;
Best Local Similarity 70.0%; Pred. No. 1.6;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FMKKNHEEV 10
:|||||:
Db 219 YLKNHEEM 228

RESULT 15

A25145
keratin, 47K type I cytoskeletal, larval - African clawed frog
N;Alternate names: XK81
C;Species: Xenopus laevis (African clawed frog)
C;Date: 22-Jul-1987 #sequence_revision 22-Jul-1987 #text_change 09-Jul-2004
C;Accession: A25145
R;Jonas, E.; Sargent, T.D.; Dawid, I.B.
Proc. Natl. Acad. Sci. U.S.A. 82, 5413-5417, 1985
A;Reference number: A25145; MUID:85270534; PMID:2410923
A;Accession: A25145
A;Molecule type: mRNA
A;Residues: 1-429 <JON>
A;Cross-references: UNIPROT:P08777; GB:M11940; NID:g214565; PIDN:AAA49894.1; PID:g214566
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 83.3%; Score 45; DB 2; Length 429;
Best Local Similarity 70.0%; Pred. No. 1.7;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FMKKNHEEV 10
:|||||:
Db 207 YLKNHEEM 216

Search completed: December 14, 2004, 09:20:21
Job time : 1.52838 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 08:45:45 ; Search time 1.96507 Seconds
(without alignments)
337.485 Million cell updates/sec

Title: US-10-026-001-7
Perfect score: 54
Sequence: 1 FMKXNHEEV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	54	100.0	232	4	US-09-919-497-78
2	54	100.0	430	4	US-09-919-039-105
3	47	87.0	400	1	US-07-730-953-2
4	45	83.3	425	4	US-09-538-092-1366
5	45	83.3	432	2	US-08-705-660-18
6	45	83.3	432	3	US-08-989-045-18
7	45	83.3	432	4	US-09-919-172-9
8	45	83.3	456	4	US-09-919-172-31
9	45	83.3	471	4	US-09-538-092-837
10	45	83.3	593	4	US-09-538-092-919
11	43	79.6	915	4	US-09-538-092-863
12	41	75.9	160	2	US-08-726-306A-183
13	41	75.9	465	4	US-09-538-092-877
14	41	75.9	466	4	US-09-610-401-3
15	41	75.9	466	4	US-09-167-206-12
16	41	75.9	469	4	US-09-538-092-948
17	39	72.2	75	4	US-09-621-976-7117
18	39	72.2	100	4	US-09-621-976-7112
19	39	72.2	106	2	US-08-893-042-1
20	39	72.2	106	4	US-09-621-976-7114
21	39	72.2	106	4	US-09-621-976-7115
22	38	70.4	466	4	US-09-610-401-4
23	37	68.5	422	2	US-09-067-351-3
24	37	68.5	422	3	US-09-360-490-3
25	36	66.7	107	2	US-08-893-042-3
26	36	66.7	343	1	US-08-454-196-6
27	36	66.7	343	2	US-08-286-819A-4

28	36	66.7	343	3	US-08-980-357-4	Sequence 4, Appli
29	36	66.7	343	3	US-09-064-033-6	Sequence 6, Appli
30	36	66.7	343	4	US-09-291-046-6	Sequence 6, Appli
31	36	66.7	715	4	US-09-134-000C-5094	Sequence 5094, Ap
32	36	66.7	2296	2	US-08-286-819A-27	Sequence 27, Appl
33	36	66.7	2296	3	US-08-980-357-27	Sequence 27, Appl
34	35	64.8	463	3	US-08-845-258-25	Sequence 25, Appl
35	35	64.8	463	3	US-08-990-571-25	Sequence 25, Appl
36	35	64.8	463	3	US-08-723-142A-25	Sequence 25, Appl
37	35	64.8	463	4	US-09-528-784A-25	Sequence 25, Appl
38	35	64.8	463	4	US-09-569-098A-25	Sequence 25, Appl
39	35	64.8	519	4	US-09-248-796A-14534	Sequence 14534, A
40	35	64.8	571	4	US-09-538-092-422	Sequence 422, App
41	34	63.0	47	4	US-09-464-152A-2	Sequence 2, Appli
42	34	63.0	109	2	US-08-893-042-4	Sequence 4, Appli
43	34	63.0	236	4	US-09-583-110-2789	Sequence 2789, Ap
44	34	63.0	258	4	US-09-328-352-4425	Sequence 4425, Ap
45	34	63.0	317	2	US-08-726-306A-168	Sequence 168, App

ALIGNMENTS

RESULT 1
US-09-919-497-78
; Sequence 78, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 78
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-78

Query Match 100.0%; Score 54; DB 4; Length 232;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKXNHEEV 10
Db 14 FMKXNHEEV 23

RESULT 2
US-09-919-039-105
; Sequence 105, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 105
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1272969CD1

US-09-919-039-105

Query Match 100.0%; Score 54; DB 4; Length 430;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
|||
Db 212 FMKKNHEEV 221

RESULT 3

US-07-730-953-2
; Sequence 2, Application US/07730953
; Patent No. 528614
; GENERAL INFORMATION:
; APPLICANT: BODENMULLER, Heinz
; APPLICANT: DESSAUER, Andreas
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF MALIGNANT
; TITLE OF INVENTION: DISEASES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaïdo, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/730,953
FILING DATE: 19910723
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 39 42 999.7
FILING DATE: 21-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kitts, Monica C.
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P564-1119
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-730-953-2

Query Match 87.0%; Score 47; DB 1; Length 400;
Best Local Similarity 70.0%; Pred. No. 1.3;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
::|
Db 213 YLKNHEEI 222

RESULT 4

US-09-538-092-1366
; Sequence 1366, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loïc
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542

US-09-538-092-1366
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1366
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q92764

US-09-538-092-1366
Query Match 83.3%; Score 45; DB 4; Length 425;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MKKNHEEV 10
:|
Db 201 LKKNHEEV 209

US-09-538-092-1366
; Sequence 18, Application US/08705660
; Patent No. 5858683
; GENERAL INFORMATION:
; APPLICANT: KEESE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,660
FILING DATE:
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-705-660-18

Query Match 83.3%; Score 45; DB 2; Length 432;
Best Local Similarity 70.0%; Pred. No. 3.2;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 83.3%; Score 45; DB 2; Length 432;
Best Local Similarity 70.0%; Pred. No. 3.2;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

US-08-705-660-18
; Sequence 18, Application US/08705660
; Patent No. 5858683
; GENERAL INFORMATION:
; APPLICANT: KEESE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,660
FILING DATE:
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-705-660-18

Query Match 83.3%; Score 45; DB 2; Length 432;
Best Local Similarity 70.0%; Pred. No. 3.2;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

US-08-705-660-18
; Sequence 18, Application US/08705660
; Patent No. 5858683
; GENERAL INFORMATION:
; APPLICANT: KEESE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,660
FILING DATE:
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-705-660-18

Query Match 83.3%; Score 45; DB 2; Length 432;
Best Local Similarity 70.0%; Pred. No. 3.2;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKHHEEV 10
:|||||:
Db 217 YLKKHHEEM 226

RESULT 6

US-08-989-045-18
; Sequence 18, Application US/08989045
; Patent No. 6027905
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,045
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-989-045-18

Query Match 83.3%; Score 45; DB 3; Length 432;
Best Local Similarity 70.0%; Pred. No. 3.2;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKHHEEV 10
:|||||:
Db 217 YLKKHHEEM 226

RESULT 7

US-09-919-172-9
; Sequence 9, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program

; SEQ ID NO 9
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 1454852CD1
US-09-919-172-9

Query Match 83.3%; Score 45; DB 4; Length 432;
Best Local Similarity 70.0%; Pred. No. 3.2;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKHHEEV 10
:|||||:
Db 217 YLKKHHEEM 226

RESULT 8

US-09-919-172-31
; Sequence 31, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 31
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 1709387CD1
US-09-919-172-31

Query Match 83.3%; Score 45; DB 4; Length 456;
Best Local Similarity 70.0%; Pred. No. 3.4;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKHHEEV 10
:|||||:
Db 239 YLKKHHEEM 248

RESULT 9

US-09-538-092-837
; Sequence 837, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 837
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P02533
US-09-538-092-837

Query Match 83.3%; Score 45; DB 4; Length 471;
Best Local Similarity 70.0%; Pred. No. 3.5;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
Db 247 YLKNHEEM 256

RESULT 10
US-09-538-092-919
; Sequence 919, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 919
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P13645
US-09-538-092-919

Query Match 83.3%; Score 45; DB 4; Length 593;
Best Local Similarity 70.0%; Pred. No. 4.3;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
Db 282 YLKNHEEM 291

RESULT 11
US-09-538-092-863
; Sequence 863, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 863
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P07197
US-09-538-092-863

Query Match 79.6%; Score 43; DB 4; Length 915;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
Db 232 FLRSNHEEV 241

RESULT 12
US-08-726-306A-183
; Sequence 183, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,049-A (3255/00784)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 183:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 160 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-726-306A-183

Query Match 75.9%; Score 41; DB 2; Length 160;
Best Local Similarity 70.0%; Pred. No. 6;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
Db 8 FLKKVHEEI 17

RESULT 13
US-09-538-092-877
; Sequence 877, Application US/09538092
; Patent No. 6753314

```
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 877
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P08670
US-09-538-092-877
```

```
Query Match 75.9%; Score 41; DB 4; Length 465;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10
|:|:|:|:|:|:|:
Db 232 FLKKLHEEEI 241
```

```
RESULT 14
US-09-610-401-3
; Sequence 3, Application US/09610401
; Patent No. 6417336
; GENERAL INFORMATION:
; APPLICANT: MORISHIMA, No. 6417336uhiro,
; APPLICANT: NAKANISHI, Keiko,
; APPLICANT: SHIBATA, Takehiko
; TITLE OF INVENTION: Antibody against cleavage product of vimentin
; FILE REFERENCE: 522.1004
; CURRENT APPLICATION NUMBER: US/09/610,401
; CURRENT FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: JP 11-193235
; PRIOR FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-610-401-3
```

```
Query Match 75.9%; Score 41; DB 4; Length 466;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10
|:|:|:|:|:|:|:
Db 233 FLKKLHEEEI 242
```

```
RESULT 15
US-09-167-206-12
; Sequence 12, Application US/09167206A
; Patent No. 6476193
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, MeiJa
; TITLE OF INVENTION: Nlk1 PROTEIN AND Nlk1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 Nlk1 protein complexes
```

```
; CURRENT APPLICATION NUMBER: US/09/167,206A
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-206-12

Query Match 75.9%; Score 41; DB 4; Length 466;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FMKKNHHEEV 10
|:|:|:|:|:|:|:
Db 233 FLKKLHEEEI 242
```

```
Search completed: December 14, 2004, 09:21:58
Job time : 2.96507 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 09:07:06 ; Search time 6.50655 Seconds
(without alignments)
548.952 Million cell updates/sec

Title: US-10-026-001-7
Perfect score: 54
Sequence: 1 FMKKNHEEV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	10	13	US-10-026-001-7
2	54	100.0	35	9	US-09-864-761-42517
3	54	100.0	122	15	US-10-264-049-3395
4	54	100.0	135	9	US-09-925-301-1529
5	54	100.0	139	14	US-10-029-386-28256
6	54	100.0	231	15	US-10-264-049-3291
7	54	100.0	232	9	US-09-919-497-78
8	54	100.0	259	15	US-10-108-260A-4615
9	54	100.0	375	15	US-10-080-334-268
10	54	100.0	423	15	US-10-080-334-269
11	54	100.0	424	15	US-10-080-334-266
12	54	100.0	430	10	US-09-919-039-105
13	54	100.0	430	10	US-09-813-432-36

14	54	100.0	430	13	US-10-026-001-1	Sequence 1, Appli
15	54	100.0	430	14	US-10-060-036-155	Sequence 155, App
16	54	100.0	430	14	US-10-174-364-36	Sequence 36, Appl
17	54	100.0	430	15	US-10-080-334-265	Sequence 265, App
18	54	100.0	430	15	US-10-080-334-267	Sequence 267, App
19	54	100.0	430	15	US-10-246-583-36	Sequence 36, Appl
20	54	100.0	430	16	US-10-689-832-36	Sequence 36, Appl
21	54	100.0	430	16	US-10-734-564-117	Sequence 117, App
22	54	100.0	432	14	US-10-106-698-5102	Sequence 5102, Ap
23	51	94.4	312	10	US-09-976-782-32	Sequence 32, Appl
24	51	94.4	312	15	US-10-080-334-167	Sequence 167, App
25	51	94.4	312	15	US-10-080-334-188	Sequence 188, App
26	51	94.4	312	15	US-10-080-334-270	Sequence 270, App
27	51	94.4	312	15	US-10-231-913-260	Sequence 260, App
28	51	94.4	542	10	US-09-976-782-4	Sequence 4, Appli
29	48	88.9	355	14	US-10-104-047-2943	Sequence 2943, Ap
30	48	88.9	402	15	US-10-080-334-88	Sequence 88, Appl
31	48	88.9	427	10	US-09-813-432-8	Sequence 8, Appli
32	48	88.9	427	14	US-10-174-364-8	Sequence 8, Appli
33	48	88.9	427	15	US-10-246-583-8	Sequence 8, Appli
34	48	88.9	427	16	US-10-689-832-8	Sequence 8, Appli
35	48	88.9	458	14	US-10-205-823-216	Sequence 216, App
36	48	88.9	483	14	US-10-316-253-8	Sequence 8, Appli
37	47	87.0	143	9	US-09-925-301-1456	Sequence 1456, Ap
38	47	87.0	180	15	US-10-264-049-3166	Sequence 3166, Ap
39	47	87.0	256	10	US-09-997-003-36	Sequence 36, Appl
40	47	87.0	400	9	US-09-922-217-1115	Sequence 1115, Ap
41	47	87.0	400	13	US-10-025-380-1115	Sequence 1115, Ap
42	47	87.0	400	16	US-10-734-564-118	Sequence 118, App
43	47	87.0	400	17	US-10-733-969A-33	Sequence 33, Appl
44	46	85.2	443	16	US-10-408-765A-2578	Sequence 2578, Ap
45	45	83.3	99	17	US-10-425-115-314778	Sequence 314778,

ALIGNMENTS

RESULT 1
US-10-026-001-7
; Sequence 7, Application US/10026001
; Publication No. US20020122791A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: ANTIGENIC CK-18 COMPOUNDS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS AND METHODS FOR USING SAME
; FILE REFERENCE: GZ 210800
; CURRENT APPLICATION NUMBER: US/10/026,001
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,820
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-001-7

Query Match 100.0%; Score 54; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FMKKNHEEV 10
Db 1 FMKKNHEEV 10

RESULT 2
US-09-864-761-42517
; Sequence 42517, Application US/03864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.

```
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42517
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL58219.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 27
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 14
; OTHER INFORMATION: SWISSPROT HIT: P05783, EVALUE 3.00e-13
; OTHER INFORMATION: EST_HUMAN HIT: T87390.1, EVALUE 3.00e-12
; US-09-864-761-42517
```

```
Query Match 100.0%; Score 54; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FMKKNHEEEV 10
Db 25 FMKKNHEEEV 34
```

```
RESULT 3
US-10-264-049-3395
; Sequence 3395, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3395
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-3395
```

```
Query Match 100.0%; Score 54; DB 15; Length 122;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FMKKNHEEEV 10
Db 44 FMKKNHEEEV 53
```

```
RESULT 4
US-09-925-301-1529
; Sequence 1529, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1529
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1529
```

```
Query Match 100.0%; Score 54; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FMKKNHEEEV 10
Db 35 FMKKNHEEEV 44
```

```
RESULT 5
US-10-029-386-28256
; Sequence 28256, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
```

```
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 28256
/ LENGTH: 139
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR22 180.0
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.55
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.6
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.8
/ OTHER INFORMATION: SWISSPROT HIT: P05783, EVALUATION 2.00e-58
US-10-029-386-28256
```

```
Query Match 100.0%; Score 54; DB 14; Length 139;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FMKKNHHEEV 10
Db 115 FMKKNHHEEV 124
```

RESULT 6

```
US-10-264-049-3291
/ Sequence 3291, Application US/10264049
/ Publication No. US20040005579A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Birse et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PA133PI
/ CURRENT APPLICATION NUMBER: US/10/264,049
/ CURRENT FILING DATE: 2002-10-04
/ PRIOR APPLICATION NUMBER: PCT/US01/18569
/ PRIOR FILING DATE: 2001-06-07
/ PRIOR APPLICATION NUMBER: US 60/209,467
/ PRIOR FILING DATE: 2000-06-07
/ NUMBER OF SEQ ID NOS: 4360
/ SOFTWARE: PatentIn ver. 3.1
/ SEQ ID NO 3291
/ LENGTH: 231
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (204)
/ OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3291
```

```
Query Match 100.0%; Score 54; DB 15; Length 231;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FMKKNHHEEV 10
Db 68 FMKKNHHEEV 77
```

RESULT 7

```
US-09-919-497-78
/ Sequence 78, Application US/09919497
/ Patent No. US2002010662A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Mutter, George L.
/ TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
/ FILE REFERENCE: B0801/7225
/ CURRENT APPLICATION NUMBER: US/09/919,497
/ CURRENT FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: US 60/221,735
/ PRIOR FILING DATE: 2000-07-31
```

```
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 78
/ LENGTH: 232
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-919-497-78
```

```
Query Match 100.0%; Score 54; DB 9; Length 232;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FMKKNHHEEV 10
Db 14 FMKKNHHEEV 23
```

RESULT 8

```
US-10-108-260A-4615
/ Sequence 4615, Application US/10108260A
/ Publication No. US20040005560A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: No. US20040005560A1 full length cDNA
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/10/108,260A
/ CURRENT FILING DATE: 2002-03-27
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4615
/ LENGTH: 259
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-108-260A-4615
```

```
Query Match 100.0%; Score 54; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FMKKNHHEEV 10
Db 247 FMKKNHHEEV 256
```

RESULT 9

```
US-10-080-334-268
/ Sequence 268, Application US/10080334
/ Publication No. US20040002584A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Pena, Carol E. A.
/ APPLICANT: Shinkets, Richard A
/ APPLICANT: Li, Li
/ APPLICANT: Shency, Suresh G
/ APPLICANT: Kekuda, Ramesh
/ APPLICANT: Spytek, Kimberly A
/ APPLICANT: Vernet, Corine A.M.
/ APPLICANT: Malyankar, Uriel M
/ APPLICANT: Guo, Xiaojia
/ APPLICANT: Gusev, Vladimir Y
/ APPLICANT: Casman, Stacie J
/ APPLICANT: Boldog, Ferenc L
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Tchernev, Velizar T
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Gangolli, Bsha A
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Baumgartner, Jason C.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Zerhusen, Bryan D
/ TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
/ USING THE SAME
```


FILE REFERENCE: 21402-275
CURRENT APPLICATION NUMBER: US/10/080,334
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/330,307
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/278,796
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/281,521
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/311,595
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/270,220
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/274,295
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/318,526
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 268
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
US-10-080-334-268

Query Match 100.0%; Score 54; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
DB 157 FMKKNHEEV 166

RESULT 10
US-10-080-334-269
Sequence 269, Application US/10080334
Publication No. US20040002584A1
GENERAL INFORMATION:
APPLICANT: Shimkets, Richard A.
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A.
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Guo, Xiaojia
APPLICANT: Gusev, Vladimir Y
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Furtak, Katarzyna
APPLICANT: Tchernev, Velizar T
APPLICANT: Patturajan, Meera
APPLICANT: Gangolli, Esha A

APPLICANT: Padigaru, Muralidhara
APPLICANT: Liu, Xiaohong
APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zerhusen, Bryan D
TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
FILE REFERENCE: 21402-275
CURRENT APPLICATION NUMBER: US/10/080,334
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/330,307
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/278,796
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/281,521
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/276,677
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/311,595
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/270,220
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/274,295
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/318,526
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 269
LENGTH: 423
TYPE: PRT
ORGANISM: Mus musculus
US-10-080-334-269

Query Match 100.0%; Score 54; DB 15; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHEEV 10
DB 205 FMKKNHEEV 214

RESULT 11
US-10-080-334-266
Sequence 266, Application US/10080334
Publication No. US20040002584A1
GENERAL INFORMATION:
APPLICANT: Pena, Carol E. A.
APPLICANT: Shimkets, Richard A
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M

```

; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 266
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-266

Query Match      100.0%; Score 54; DB 15; Length 424;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FMKKNHEEV 10
Db      206 FMKKNHEEV 215

RESULT 12
US-09-919-039-105
; Sequence 105, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 105
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1272969CD1
US-09-919-039-105

Query Match      100.0%; Score 54; DB 10; Length 430;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FMKKNHEEV 10
Db      212 FMKKNHEEV 221

RESULT 13
US-09-813-432-36
; Sequence 36, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Majmuder, Kamud
; APPLICANT: Spaderna, Steven K
; APPLICANT: Smithson, Glenda
; APPLICANT: Mezes, Peter S
; APPLICANT: Vernet, Corine A. M.
; TITLE OF INVENTION: No. US20030148485A1 Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729
; CURRENT APPLICATION NUMBER: US/09/813,432
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-432-36

Query Match      100.0%; Score 54; DB 10; Length 430;
```

Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHREEV 10
Db 212 FMKKNHREEV 221

RESULT 14

US-10-026-001-1
; Sequence 1, Application US/10026001
; Publication No. US20020122791A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: ANTIGENIC CK-18 COMPOUNDS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS AND METHODS FOR USING SAME
; FILE REFERENCE: GZ 210800
; CURRENT APPLICATION NUMBER: US/10/026,001
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,820
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-001-1

Query Match 100.0%; Score 54; DB 13; Length 430;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHREEV 10
Db 212 FMKKNHREEV 221

RESULT 15

US-10-060-036-155
; Sequence 155, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 155
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-155

Query Match 100.0%; Score 54; DB 14; Length 430;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FMKKNHREEV 10
Db 212 FMKKNHREEV 221

Search completed: December 14, 2004, 09:27:03
Job time : 7.50655 secs